Title
Teaching For All? Variation in the Effects of Teach For America

Author
Penner, Emily Kathryn

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Teaching For All? Variation in the Effects of Teach For America

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Education

by

Emily Kathryn Penner

Dissertation Committee:
Distinguished Professor Greg J. Duncan, Chair
Founding Dean and Professor Deborah Lowe Vandell
Professor George Farkas
Professor Marianne P. Bitler
Associate Professor Thurston Domina

2014

DEDICATION

To
los estudiantes del Salón 16

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank Mathematica Policy Research for making data available and Steve Glazerman for his assistance with data questions. I would also like to thank the North Carolina Education Research Data Center and the External Research Partnerships team at Teach For America for making data available and for their invaluable assistance with matching Teach For America teachers to administrative teacher records.

I am grateful to Greg J. Duncan, Deborah Lowe Vandell, George Farkas, Marianne Bitler, Thurston Domina, Andrew Penner, and Anna Penner for their valuable comments. I would also like to thank seminar participants at the School of Education and the Population, Society, and Inequality Colloquium Institute at the University of California, Irvine, the Department of Social Science and Policy at New York University, Abu Dhabi, the Department of Educational Psychology in the College of Education at the University of Illinois, Urbana-Champaign, the School of Education at the University of California, Davis, and the School of Education at the University of Michigan for their feedback and suggestions.

I would also like to thank my family for their support during this program, particularly Andrew Penner, Kay and Wendell Wolff, Ken Gjeltema and Cyndi Vasallo, Mark and Mary Esther Penner, Trevor Gjeltema and Heather Woodford, Anna Penner, Daniel Penner, Aaron Wolff, Dane Wolff, and Hannah Wolff, and my Irvine family, the Yis, the Beenens, the Ormonds, the Strodtbecks, the Weltons, and the Andos. Finally, I would like to thank C3 for their encouragement and friendship.

Research reported in this dissertation 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.

CURRICULUM VITAE

EMILY K. PENNER

University of California–Irvine

EDUCATION

Ph.D. in Education, Specialization in Education Policy and Social Context — 2014
University of California, Irvine

Dissertation: Teaching For All? Variation in the Effects of Teach for America
Committee: Greg J. Duncan (Chair), Marianne P. Bitler (Economics), Thurston Domina, George Farkas, Deborah Lowe Vandell

M.A. in Education — 2012
University of California, Irvine

M.A. in Education, with Multiple Subject Teaching Credential — 2006
Alliant International University

B.A. in Economics and International Relations, with Honors — 2005
Claremont McKenna College
cum laude

RESEARCH INTERESTS

Inequality, education policy, teacher effectiveness, parenting, quantitative methods, and distributional methods.

PEER REVIEWED PUBLICATIONS

Domina, Thurston, Andrew J. McEachin, Andrew M. Penner, Emily K. Penner, AnneMarie Conley. "Aiming High and Falling Short: California's 8th Grade Algebra-For-All Effort." *Educational Evaluation and Policy Analysis*. Forthcoming.

Bitler, Marianne P., Thurston Domina, Hilary Hoynes, and Emily K. Penner. "Distributional Effects of a School Voucher Program: Evidence from New York City." NBER Working Paper No. w19271. *Journal of Research on Educational Effectiveness*. Forthcoming.

Domina, Thurston, Andrew M. Penner, Emily K. Penner, AnneMarie Conley. "Algebra for All: California's 8th grade Algebra Initiative as Constrained Curricula." *Teachers College Record.* Forthcoming.

Auger, Anamarie, Stephanie M. Reich, and Emily K. Penner. 2014. "The Effect of Baby Books on Mothers' Reading Beliefs and Reading Practices." *Journal of Applied Developmental Psychology, 35:* 337-346. DOI: 10.1016/j.appdev.2014.05.007

Penner, Emily K. 2012. "Education Policy and Conflict in Latin America: Lessons from Chile and Venezuela." Pp. 181-212 in *Economic Development Strategies and the Evolution of Violence in Latin America* (William Ascher and Natalia Mirovitskaya, Eds.) New York, New York: Palgrave Macmillan.

Reich, Stephanie M., Emily K. Penner, Greg J. Duncan, and Anamarie Auger. 2012. "Using Baby Books to Change New Mothers' Attitudes About Corporal Punishment." C*hild Abuse & Neglect, 36(2): 108-117.*

Reich, Stephanie M., Emily K. Penner, and Greg J. Duncan. 2011. "Using Baby Books to Increase New Mothers' Safety Practices." *Academic Pediatrics, 11(1): 34-43.*

OTHER PUBLICATIONS

Bitler, Marianne, Thurston Domina, Andrew M. Penner, and Emily K. Penner. Forthcoming. "Quantile Regression." Forthcoming in Dominic J. Brewer and Lawrence O. Picus (Eds.), *Encyclopedia of Education Economics and Finance*. Thousand Oaks, CA: Sage.

WORK UNDER REVIEW

Penner, Emily K. "Teaching For All? Teach For America's Effects on the Distribution of Student Achievement."

Penner, Emily K. "Parenting and the Reduction of Inequality: How the Impact of Early Parenting on Achievement Varies across Class and with Development."

Thurston Domina, Andrew M. Penner, and Emily K Penner. "Social Rewards: Incentives, Stigma, and Stereotype Threat in High School."

Penner, Andrew M., Thurston Domina, Emily K. Penner, and AnneMarie Conley. "The Distributional Effects of Curricular Intensification."

WORK IN PROGRESS

Bitler, Marianne P., Sean P. Corcoran, Thurston Domina, and Emily K. Penner. "Teacher Value-Added and Noise: An Analysis of Teacher Effects on Achievement and Anthropometrics."

Penner, Emily K. "Teaching For America? The Good, the Bad, and the Unknown."

HONORS AND AWARDS

Society for Research in Child Development Dissertation Funding Award	2013
UC Irvine Department of Education Travel Awards	2010-2012
UCEC Innovative Evaluation Research Award	2011
Invited Participant, UCEC Institute for Training in Educational Evaluation	2011
International Relations Outstanding Student Award	2005
Baker Lowe Political Economy Scholar	2004-2005
President of Claremont International Relations Society	2004-2005

INVITED PRESENTATIONS

Penner, Emily K. "Teaching For All? Variation in the Effects of Teach For America." Presented at:
Population, Society and Inequality Brown Bag, Department of Sociology. University of California, Irvine. November 2013.
Department of Social Science and Policy. New York University, Abu Dhabi. November 2013.
Department of Educational Psychology, College of Education. University of Illinois, Urbana-Champaign. December 2013.
School of Education. University of California, Davis. February 2014.
School of Education. University of California, Irvine. February 2014.
School of Education. The University of Michigan. March 2014.

Penner, Emily K. "Education Policy and Conflict in Latin America: Lessons from Chile and Venezuela." Presented at the Conference on Politics, Economics, and Inclusive Development, Pacific Basin Research Center, SOKA University of America, Aliso Viejo, California, November 2011.

PEER REVIEWED CONFERENCE PRESENTATIONS

Thurston Domina, Andrew M. Penner, and Emily K Penner. "Social Rewards: Incentives, Stigma, and Stereotype Threat in High School," Presented at the XVIII International Sociological Association, World Congress of Sociology, Yokohama, Japan, July, 2014.

Penner, Emily K. "Does Teach For America Have Long-Term Impacts?" Presented at the Annual Meeting of the American Educational Research Association, Philadelphia, PA, April, 2014.

Penner, Emily K. "Does Teach For America Have Long-Term Impacts?" Presented at the Spring Conference of the Society for Research on Educational Effectiveness, Washington D.C., March, 2014.

Penner, Emily K. "Does Teach For America Have Long-Term Impacts?" Presented at the Annual Meeting of the Sociology of Education Association, Pacific Grove, California, February, 2014.

Penner, Emily K. "Treatment Effect Heterogeneity in Teach For America." Presented at the Annual Meeting of the American Educational Research Association, San Francisco, California, April, 2013.

Penner, Emily K. “Missing Data and Mixed Results: The Effects of Teach For America on Student Achievement Revisited.” Presented at the Spring Conference of the Society for Research on Educational Effectiveness, Washington D.C., March, 2013.

Penner, Emily K. “Treatment Heterogeneity in the Effect of Teach For America.” Presented at the Annual Meeting of the Sociology of Education Association, Pacific Grove, California, February, 2013.

Domina, Thurston, Emily K. Penner, Andrew M. Penner, and AnneMarie Conley. “Does Detracking Work? Evidence from a Mathematics Curricular Reform.” Presented at the Fall Conference of the Association for Public Policy Analysis and Management, Baltimore, Maryland, October, 2012.

Penner, Emily K. “The Equitable Effects of Early Concerted Cultivation on Elementary and Middle School Achievement.” Presented at the Annual Meeting of the American Educational Research Association, Vancouver, British Columbia, Canada, April, 2012.

Bitler, Marianne, Thurston Domina, and Emily K. Penner. “Teacher Effects on Achievement and Anthropometrics: A Cautionary Tale.” Presented at the Annual Meeting of the American Educational Research Association, Vancouver, British Columbia, Canada, April, 2012.

Domina, Thurston, Emily K. Penner, Andrew M. Penner, and AnneMarie Conley. “Detracking Across the Distribution: Evidence from a Mathematics Curricular Reform.” Presented at the Spring Conference of the Society for Research on Educational Effectiveness, Washington D.C., March, 2012.

Bitler, Marianne, Thurston Domina, Hilary Hoynes, and Emily K. Penner. “Distributional Effects of a School Voucher Program: Evidence from New York City.” Presented at the Spring Conference of the Society for Research on Educational Effectiveness, Washington D.C., March, 2012.

Penner, Emily K. “Parenting and the Reduction of Inequality: How the Impact of Early Parenting on Achievement Varies Across Class Boundaries and with Development.” Presented at the Annual Meeting of the Sociology of Education Association, Pacific Grove, California, February, 2012.

Domina, Thurston, Andrew M. Penner, Emily K. Penner, and AnneMarie Conley. “Does Detracking Work? Evidence from a Mathematics Curricular Reform.” Presented at the Annual Meeting of the American Sociological Association, Las Vegas, Nevada, August, 2011.

Penner, Emily K.. “Who Does Concerted Cultivation Help? Examining the Distributional Effects of Concerted Cultivation.” Presented at the Annual Meeting of the American Sociological Association, Las Vegas, Nevada, August, 2011.

Reich, Stephanie M., Leonard Bickman, Greg Duncan, Emily K. Penner, Anamarie Auger, Benjamin Saville, JoAnn Alvarez, Ana Regina Andrade, and Alejandra Albarran. "Using Baby Books to Increase Knowledge, Improve Maternal Attitudes, Beliefs, and Practices, and Promote Child Health." Presented at the Meeting of the Pediatric Academic Society, Denver, Colorado, May, 2011.

Auger, Anamarie, Stephanie M. Reich, Emily K. Penner, and Greg J. Duncan. "The Effect of Baby Books on Mothers' Reading Beliefs and Reading Practices." Presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada, April, 2011.

Domina, Thurston, Emily K. Penner, Andrew M. Penner, and AnneMarie Conley. "Detracking Across the Distribution: Evidence from a Mathematics Curricular Reform." Presented at the Annual Meeting of the Sociology of Education Association, Pacific Grove, California, February, 2011.

RESEARCH EXPERIENCE

Graduate Student Researcher, Marianne P. Bitler 2011-Present
NICHD P-01 P01HD065704
"Human Capital Interventions across Childhood and Adolescence"
Program Director: Greg J. Duncan
Sub-project III: Distributional Impacts of School Policies
Responsibilities: Analyzed restricted-use longitudinal data from a voucher experiment using standard quantile regression, instrumental variables quantile regression, and inverse-propensity score weighted quantile regression; analyzed ECLS-K data using fixed-effects regression models; prepared tables and figures and drafted manuscripts; presented findings at national conferences and the annual advisory board meetings of the Irvine Network for Interventions in Development; investigated new projects and data sources; developed research partnerships with national organizations and off-campus collaborators.

Graduate Student Researcher, Thurston Domina 2010-Present
UC Educational Evaluation Center (UCEC) & NICHD P-01
Responsibilities: Assembled, cleaned, and analyzed administrative data from three Southern California school districts; analyzed ECLS-K data using fixed and random effect value-added models; drafted and revised manuscripts and prepared tables and figures; presented findings at national conferences; assisted in the preparation of grant proposals.

Graduate Student Researcher, Greg J. Duncan 2010-2011
William T. Grant Foundation
"Replication and Extension of a Study of Peer Impacts on Attitudes and Drinking Behavior"
UC Irvine Principal Investigator: Greg J. Duncan

Responsibilities: Cleaned, constructed, and analyzed random assignment roommate data using fixed effects regression models; analyzed ECLS-K data using fixed effects regression; assisted in original data collection and coding.

Graduate Student Researcher, Stephanie Reich 2009-2010
NICHD, RO1HD047749-01A1
"Baby Books Project: Using Baby Books to Promote Maternal and Child Health"
UC Irvine Principal Investigator: Stephanie Reich
Responsibilities: Cleaned, constructed, and analyzed longitudinal dataset; created tables and contributed to drafting and revising of manuscripts.

Research Assistant, William Ascher 2004-2005
Responsibilities: Analyzed state and city ballot initiatives and environmental policy changes, drafted policy memos survey questions.

Student Researcher, University of Havana, Havana, Cuba 2002
Responsibilities: Developed independent research project examine household remittance use: designed questionnaires, planned sampling scheme, interviewed 30 remittance recipients in money transfer centers throughout Havana in Spanish, and analyzed interview data.

GRADUATE TEACHING EXPERIENCE

Guest Lecture, School of Education 287: Data Analysis in Education 2013
School of Education, UC Irvine
Lecture for first through third year graduate students on Data & File Management for Reproducibility with STATA

Workshop on Designing Research with the ECLS-K 2011-2013
School of Education, UC Irvine
Interactive graduate student workshop on planning individual class and PhD milestone projects using the ECLS-K (given once each year)

K-12 TEACHING EXPERIENCE

Literacy and ELD Coach, Hannalei Elementary School 2008-2009
Lindamood Bell Clinician, Hannalei Elementary School 2008-2009
Founding Teacher and Grade Level Chair, Greenleaf Elementary School 2007-2008
Bilingual and Structured English Immersion Teacher, Whittier Elementary 2005-2007
Teach for America Corps Member, Bay Area '05
Placement: Bilingual first and second grade 2005-2007

SERVICE

Reviewer: *Journal of Policy Analysis and Management, Sociology of Education, Social Science Research*
Cohort Representative to the Associated Doctoral Students in Education 2012-2013

Ph.D. Student Recruitment Committee 2011

SKILLS AND CERTIFICATIONS

STATA, with training in longitudinal data analysis, advanced regression models (including fixed effects, propensity score matching, regression discontinuity, and instrumental variables), quantile regression, value-added models, HLM, and multiple imputation.
AMOS, with training in Structural Equation Modeling
Fluent Spanish
Lindamood-Bell reading fluency and comprehension intervention certified
California Multiple Subjects Credential (cleared)
Spanish Bilingual, Cross-cultural, Language, and Academic Development Credential (BCLAD)

ABSTRACT OF THE DISSERTATION

Teaching for All? Variation in the Effects of Teach For America

By

Emily Kathryn Penner

Doctor of Philosophy in Education

University of California, Irvine, 2014

Distinguished Professor Greg J. Duncan, Chair

Teach For America (TFA) is a high-profile alternative teacher certification program that seeks to provide high-quality teachers to students in low-income schools. While a growing body of research examines the average impact of TFA, few researchers examine how its impacts may vary across students and whether it positively impacts all children.

This dissertation broadens our understanding of the impacts of TFA by examining variation in its effectiveness across a number of dimensions. The first study uses experimental data from Mathematica Policy Research to examine the effect of TFA teachers on the distribution of student achievement in elementary school. Results suggest that, relative to non-TFA teachers, TFA teachers have a positive average effect on math achievement that is shared across much of the distribution, while in reading TFA has a negative effect on the bottom of the distribution and a positive effect on the top, particularly relative to veteran non-TFA teachers.

The second and third studies use administrative data from North Carolina. The second study tests for variation in the effect of TFA across grade levels and subject areas and finds that TFA has positive effects in most subjects, which are largest in high school science and math. In addition, examinations of the relative fit between TFA and student baseline proficiency suggest

that TFA teachers have larger effects for initially higher-performing students in elementary and middle school, but that in high school their effects are larger among initially lower-achievers, underscoring the importance of considering the person-environment fit. The third study examines whether TFA's effects on student achievement have changed over time and demonstrates that TFA's efficacy has increased as the program has matured, and that this change is not accounted for by most observable measures of teacher quality.

Together, the three studies provide new evidence about the benefits and shortcomings of TFA, improving our understanding of the efficacy and potential of TFA and similar programs. These results provide policy-relevant information that speaks to larger debates about alternative certification and teacher quality in high-poverty schools, and provide evidence regarding the effective allocation and training of teachers for high-poverty schools through both alternative and traditional teacher training programs.

Introduction

Teach For America (TFA) is an influential component of today's teaching landscape, but its effects on student performance, student learning, and the field of education more broadly, are still only minimally understood. TFA's presence throughout the country expanded from 500 corps members in five regions in 1990 to more than 10,000 TFA corps members in 46 regions across 36 states and the District of Columbia serving 750,000 students in 2013 (Barahona, 2012; teachforamerica.org, 2013b). These teachers represent roughly one tenth of one percent of the 7.2 million teachers in classrooms across the country (United States Census Bureau, 2011) and, a small fraction of teachers in their placement districts, and yet TFA has had an increasing influence on education policy and the field of teaching over the past two decades. However, our understanding of what TFA does and whether TFA is improving opportunities for the students it serves is still quite limited.

With its expansion, TFA has become increasingly controversial. It has drawn substantial criticism from academics, former corps members, teachers, teachers unions, and the media for hindering student achievement, undermining education and teacher certification reform, contributing to cyclical teacher turnover in hard to staff districts, using resources that would be better spent elsewhere, and many more perceived shortcomings (Bruenig, 2012; Darling-Hammond, 2011; Miner, 2010; Schorr, 1993; Socol, 2008; Vasquez Heilig & Jez, 2010). Proponents, on the other hand, argue that it has promoted student achievement, improved teacher quality, and breathed new life and dynamism into the field of education (Amos, 2012; Eidler, 2013; Guter, 2012; Higgins, Robison, Weiner, & Hess, 2011). While the critiques and praise for

the program raise many valid points, the central question in this debate is the degree to which TFA does or does not benefit students in TFA classrooms.

In order to evaluate this program objectively, it is essential to develop a careful understanding of the organization and its teachers, and use this understanding to inform research addressing whether TFA is having its intended effect on student performance. Once we have more definitive evidence about the efficacy of the program in this core domain, broader questions about the secondary effects of TFA can be considered in light of the evidence on student learning.

Assessing TFA's effectiveness is not only relevant for evaluating whether this particular program improves student performance in underserved schools. TFA is also important because it makes a provocative argument about the nature of teacher selection and training. From its inception, TFA has challenged the notion that effective teachers need extensive training 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 be effective with minimal training or teaching experience. Even more controversial, TFA promotes the idea that teacher quality can be measured, and that high quality teachers can be identified through a rigorous selection process that seeks out several personal qualities without evidence of specific teaching skill. If TFA is at all effective in improving student achievement, this suggests that alternative pathways to teacher certification are a promising complement to existing teacher preparation programs. In contrast, if TFA has a null or negative impact, its use and proliferation should be reconsidered.

Since its inception TFA has defined its success in terms of improvements to student achievement. Early on in its existence, TFA focused on "closing the achievement gap" for

students in the schools it serves, and emphasized achieving "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 get students caught 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). In addition to changing its core mission, TFA has also revised its teacher training program, with a particular focus on measuring student learning, and expanded its recruitment and placement efforts to expand its national presence.

Despite this focus on student achievement, very little is known about how TFA teachers compare with same-school and same-district counterparts on achievement metrics. As the program evolves, the pool of evidence examining the performance of TFA on achievement-based metrics continues to improve, but it is still limited and often focuses on the average effects of TFA. The best experimental and quasi-experimental evidence suggests that TFA teachers are more effective on average than other teachers in math and science, but not in reading, at the elementary and high school levels (Clark et al., 2013; Glazerman, Mayer, & Decker, 2006; Kane, Rockoff, & Staiger, 2008; Xu, Hannaway, & Taylor, 2011). Quasi-experimental and descriptive evidence using administrative data from a variety of locations yields more mixed conclusions, with some finding positive effects of TFA teachers relative to non-TFA teachers, others finding a varied pattern by subject, and still others finding that TFA teachers performed no differently or worse than non-TFA teachers (Darling-Hammond, Brewer, Gatlin, & Vasquez Heilig, 2005; Schoeneberger, Dever, & Tingle, 2009; Ware et al., 2011).

This mixed pattern of effects suggests that TFA as an intervention might have heterogeneous impacts rather than a uniform effect for all types of students in all contexts.

Developmental science suggests that many types of interventions affect students differently (Duncan & Vandell, 2011; Weiss, Bloom, & Brock, 2013), and may depend on the fit between the students and their environments in addition to the strength of the intervention itself (Eccles & Midgley, 1989). Existing examinations of TFA have only begun to consider how its effects might vary across different types of students and settings. Much of the previous literature examining variation in the effect of TFA focused on differences in the qualifications of the teachers (see e.g. Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Kane et al., 2008). Fewer studies have focused on differential effects based on demographic characteristics of the students who experience this intervention or their prior achievement levels (but see Glazerman et al., 2006; Xu et al., 2011). While these studies highlight the idea that the effects of TFA vary, and examine its average effectiveness among demographic and prior performance student subgroups, little other work considers the possibility that the impact of TFA on student achievement could vary along other dimensions.

Importantly, TFA might have heterogeneous effects for several reasons. For example, in focusing on student achievement in attempts to help students catch up to or exceed grade-level standards, TFA attempts to intervene on behalf of students in a way that suggests distributional variation in its effects. Students enter TFA classrooms with different skills and achievement levels that put them closer or farther from meeting end-of-grade-level benchmarks. As TFA teachers work to help their students meet or exceed these benchmarks, some students will need to learn more than others, with those at the lowest part of the distribution needing to make the largest gains. If TFA teachers are meeting these goals where other teachers are not, then the effect of TFA may not be uniform across the distribution of student achievement, and we would expect the largest gains at the bottom of the distribution.

This study seeks to better understand the effects of TFA by examining how these effects vary across 1) the distribution of student achievement, 2) developmental stages and content areas, and 3) the program's evolution over time, and the extent to which these effects are explained by differences and changes in teacher quality. In doing so, these studies highlight the idea that the effects of TFA vary, and begin to lay the groundwork for a more comprehensive analysis that will enable us to understand the disparate impacts that TFA has, and inform larger conversations about teacher training and effectiveness in underserved communities.

Specifically, this study extends our knowledge about the relationship between TFA and student achievement by considering multiple dimensions of heterogeneity in the relationship between TFA and student outcomes in three empirical chapters. In the first empirical chapter, I examine the impact of TFA across the distribution of student achievement in elementary school, comparing students at various percentiles of the TFA and non-TFA distribution. In the second, I test for differences in the relationship between TFA and student achievement based on the fit between the TFA program and students' developmental stage and baseline proficiency. Finally, in the third empirical chapter, I investigate whether TFA's effects have changed as the program has matured and whether changing program effects are explained by teacher quality. Below I summarize each study in turn, and provide a brief overview of my findings.

Study 1: Teaching For All? TFA's Effects on the Distribution of Student Achievement

In this chapter, I examine the effect of TFA on the distribution of student achievement in elementary school using experimental data from Mathematical Policy Research (MPR). This experiment randomized students attending the same school and grade level to TFA and non-TFA classrooms within those schools in five TFA sites across the country for one year. I use these data to estimate Quantile Treatment Effects (QTE, the differences between the TFA and non-

TFA students at various percentiles of the post-test distribution) to examine the impact of TFA on the distribution of student achievement. The experimental nature of the data allows for causal estimates of this relationship, and eliminates bias due to sorting of students into classrooms. Distributional results reveal a positive impact of TFA teachers across the distribution of math achievement relative to non-TFA teachers in the same grade levels and schools (ES as large as 0.34 standard deviations). However, the pattern of effects is different for reading. Students at the bottom of the reading distribution scored the same or worse in TFA classrooms (ES as small as -0.19 standard deviations), while in the upper half of the distribution, students in TFA classrooms outperformed students in non-TFA classrooms (ES as large as 0.19 standard deviations). Comparisons relative to veteran non-TFA teachers confirm this pattern, but are more pronounced. This suggests that TFA teachers may not be a good fit in environments in which students require intensive basic literacy instruction and reading intervention, where traditional training and experience are instead superior. However, TFA teachers' strong math content knowledge is beneficial for students across a range of skill levels.

Study 2: Stage-Proficiency-Environment Fit and the Effectiveness of Teach For America

My second chapter builds on a growing literature examining the mean impact of TFA on student achievement by testing whether it varies by student baseline proficiency and across developmental stages in elementary, middle, and high school and across the subject areas of language arts, social studies, math, and science. Although TFA places teachers in classrooms across pre-kindergarten through twelfth grade, and provides them with training targeted towards their school level and subject specific assignment, TFA is often evaluated as a uniform intervention that is presumed to have similar impacts across these distinct developmental stages

and different school structures. Uniform impacts across school levels and subjects are possible, but they have only been tested by two studies (Henry et al., 2010; Ware et al., 2011).

To test for variation in the relationship between TFA and student achievement across developmental stages and entering skill levels, I use administrative data from the State of North Carolina for students in grades 3-12 for the school years 1999/2000 through 2010/2011, which was provided by the North Carolina Education Research Data Center (NCERDC) at Duke University. For this study, I focus only on the students and teachers in the 2005/2006 – 2010/2011 school years, once the TFA program was well established. TFA teachers from the two TFA regions in North Carolina were identified with assistance from Teach For America.

I estimate the relationship between having a TFA teacher and student achievement using OLS regression to make two types of comparisons. First, I use school-grade-year (or school-subject-year in high school) fixed effects to compare student achievement outcomes for students in TFA classrooms relative to students in non-TFA classrooms in the same school, grade (or subject), and year. These fixed effects enable me to compare TFA students relative to students who had the potential to have a TFA teacher because they attended the same school and were in the same grade or subject, but were not assigned to a TFA teacher, rather than relative to all students that were not in TFA classrooms, but did not have the potential to be in a TFA classroom. Second, I use student fixed effects to compare students' achievement in the year they had a TFA teacher to their achievement in years in which they did not have a TFA teacher, and student-year fixed effects to compare across high school subjects in which they did and did not have a TFA teacher within the same year. This allows me to use students as their own counterfactual to test whether having a TFA teacher was associated with improvements in student achievement across subjects in the same year or within subjects across years. I use these

two approaches to compare students across developmental stages, subjects, and proficiency levels.

My results suggest that TFA has significant effects on student achievement across all three developmental stages, with the largest effects found in high school. I also find evidence of variation in the effects of TFA across academic domains, as TFA teachers are most effective in math and science, particularly in high school, where TFA students outperform non-TFA students by as much as 0.20 standard deviations). I also find some evidence of marginally significant negative effects on elementary reading, a finding that echoes the results presented in Chapter 2. Finally, I find evidence that TFA effects vary based on the match between grade level and students' incoming skill levels. TFA effects in elementary and middle school are largest for students with the highest baseline test scores in both math and reading (by as much as 0.13 standard deviations in elementary reading), whereas across all high school subjects, TFA's effects are largest for students with the lowest entering test scores (which are as much as 0.08 standard deviations larger in high school social studies). Together, these findings suggest that TFA's effects are far from uniform, but are nonetheless positive in most areas, save elementary reading. These results also underscore the need to test for impact variation and to investigate how a program is affected by the fit between with the developmental stage and entering skill level of the children it hopes to serve.

Study 3: Teacher Quality and Teach For America's Increasing Effectiveness over Time

My third dissertation study examines whether TFA's effects have increased as the program has matured and whether these effects, or changes in effects, are explained by measures of observable teacher quality. The existing literature on TFA treats the program as though it has not changed over time, and likewise does not consider the possibility of changes in

counterfactual teacher quality. However, the program has expanded substantially and revised its training and mentoring program multiple times since its inception. One notable example occurred in the mid-2000s. At this time TFA embarked on a massive expansion program and began to incorporate the collection and measurement of student achievement data into its training program and internal organizational benchmarks. At the same time, North Carolina also expanded its own alternatively certified teacher corps, and other local and regional training programs may have undergone changes in response to the presence of these competing teacher training programs. Although TFA has continued to evolve to better achieve its mission and respond to critiques, Study 3 is the first place where the effects of this type of organizational change on student outcomes are evaluated.

As with Study 2, this study uses administrative data from North Carolina, but it focuses on the full panel of data from the 1999/2000 to 2010/2011 school years, rather than just the more recent years. This is the longest panel of data used in a study of TFA to date, and allows me to compare whether TFA effects have changed before and after 2005, which marked a period of expansion in North Carolina that mirrored the national expansion. I also take advantage of several measures of teacher quality, which have been recorded in North Carolina since 1995 and include items like whether the teacher is fully certified, if they have a master's degree, and what their Praxis certification scores were. I use these measures of teacher quality to test whether TFA's effects on achievement overall or changes in TFA's effects over time are explained by teacher quality. As with Study 2, I use school-grade-year or school-subject-year fixed effects to compare the effect of TFA teachers on student achievement relative to students who could have been assigned to TFA classrooms, but were not.

I find that TFA's impacts on student achievement have improved over time in several subjects and grade levels relative to non-TFA teachers, specifically middle school math, high school science, and high school social studies. This pattern of improvement is shared more broadly across several other subjects and grades, although the differences are only marginally significant. In some of those cases, TFA had a positive impact early on and its impact grew over time, as in high school science where the TFA effect is 0.09 standard deviations higher in the later time period. In others, TFA's impact was initially negative and improved substantially in later years, as in high school social studies where it improved by 0.54 standard deviations..

I also find that the main effects of TFA on student achievement are not explained by most observable teacher quality indicators. The only exception is teacher Praxis scores[1], which accounts for approximately one third of TFA's effect on student achievement across subject areas and grades. Likewise, none of the other measures of teacher quality explain TFA's improvements after 2005, except for average Praxis scores, where teachers' Praxis scores account for approximately a quarter of the improvement in TFA effects over time in the three subjects in which TFA's effects grew.

These results suggest that TFA's program has changed in important ways over time that have led to improvements in its effects on student achievement in some subjects, but not others. They also suggest that measures of observable teacher quality that are typically used in personnel decisions do little to identify the effects of TFA, which may have improved due to TFA selecting teachers for other qualities or because of improvements to the program more broadly (e.g., better training or ongoing support). This is an ongoing question that is worthy of further study. While my results do not provide evidence of positive TFA effects across the board, they suggest that

TFA teachers are better than the currently available non-TFA teachers who are employed in the same schools and subjects/grades with only one clear exception, elementary reading.

Conclusion

Finally, I conclude by highlighting five key lessons from the three substantive chapters in this dissertation. First, the effects of TFA, while typically positive, are largest in science and math and in the later grades. This suggests that the advantage of having a TFA teacher may be largest where the content being taught is more complex. Second, TFA teachers struggle to teach elementary reading, and in particular struggle teaching reading to the lowest achieving students. This suggests that the TFA model may have an important shortcoming, and that the training received may not adequately equip TFA teachers to successfully intervene and help their struggling readers. Third, TFA teachers in many subject areas and grades outperform even veteran teachers teaching the same subjects/grades in the same schools, in the same years. This indicates that TFA is largely successful at boosting student achievement in the hard-to-staff schools where they aim to improve teaching. Fourth, while the (largely) positive effects of TFA are impressive, and in the best cases are close to a third of a standard deviation, this is not a large enough effect to meet TFA's goal of closing the achievement gaps. This means that while TFA is helping to narrow achievement differences, considerable work needs to be done, both within and outside of TFA to close these gaps. Fifth, and finally, these studies highlight the importance of examining impact variation. These results can be useful for district administrators seeking to understand where TFA teachers might have the largest impacts, to TFA leadership seeking to enhance the program by addressing its weaknesses, and to other certification programs seeking to learn from the successes and failures of TFA.

[1] Praxis tests measure teacher candidates' knowledge and skills and are used for licensing and certification in North Carolina as well as many other states across the country (ETS, 2014).

Chapter 1

Literature Review

In reviewing the literature for this dissertation, I begin by presenting a broad overview of Teach For America (TFA), defining and describing the components of TFA and reviewing what we know about the impacts of TFA, providing a background for this dissertation and the larger research agenda out of which it emerges. I then focus more specifically on why TFA might have heterogeneous effects on student achievement, describing some of the more important sources of variation. Portions of this literature review are presented more concisely in the relevant empirical chapters as they would appear in articles submitted for publication.

What is TFA?

It will be useful to conceptualize TFA not as a single intervention, but rather as a system of interventions. Figure 1.1 presents a conceptual model depicting how the various elements of TFA operate, linking program components to intended outcomes.

[Insert Figure 1.1 Here]

I use this logic model to organize the existing evidence on the effects of individual components and TFA as a whole. The model parses out the components of the TFA "program" into specific components and connects them to their intended outcomes. As I describe each component, I discuss the literature and relevant findings for each component. After reviewing the research that examines these questions, I briefly address research on the effects of TFA outside of this primary structure, discussing the positive and negative externalities created by the program. I conclude by using this framework to identify areas where important questions about TFA remain,

highlighting how my dissertation begins to address these gaps in our understanding, and briefly discussing a broader research agenda to provide a more comprehensive understanding of TFA.

Understanding TFA. Teach For America is often described and evaluated simply as a program that recruits and trains teachers to teach in high need classrooms throughout the U.S. Brief mention is made of the training program and ongoing support, but most studies provide little detail about the program itself or how it is intended to have an impact. This view of TFA as a singular entity inhibits our ability to isolate those aspects of the program that make the largest contributions to its successes (and shortcomings). To the degree that TFA is a relatively permanent feature of the U.S. educational landscape, unpacking the distinct components of TFA and discussing their connections to the intended outcomes will help to deepen our understanding of how the program works, and may provide useful insights as to what parts of TFA are the most helpful and what aspects should receive the most attention for improvement.

This model centers on the founding idea of TFA, which is to improve students' educational opportunities and achievement. TFA began as Wendy Kopp's senior thesis in 1989 with the intention of addressing educational inequities and, providing all children with the "opportunity to attain an excellent education" (Kopp, 2011, p. 174). Her principal mechanism for doing so was by changing the types of individuals recruited to the teaching force to staff underserved urban and rural school districts throughout the country. Since its inception and with few modifications, TFA continues to recruit recent college graduates and professionals with strong backgrounds in academics and leadership, providing them with five weeks of intensive training and ongoing support during their two-year commitment in their placement classrooms. TFA's central goals have shifted over time. At the beginning, Kopp's project intended to build "a national teaching corps" to promote educational opportunities (Kopp, 2011) . In the late 1990s

and early 2000s, this shifted to a focus on "closing the achievement gap." TFA teachers were urged to provide their students with "significant gains," meaning 1.5 to 2 grade levels of improvement in core subject areas (Foote, 2009; Raymond, Fletcher, & Luque, 2001). It has since expanded its mandate to, "alter educational trajectories," (Kopp & Farr, 2011) which expands its mission to focus on promoting a passion for learning, and school engagement. Despite the shifting goals, TFA describes its own success in terms of the impact it has on student academics, whether that be through test scores, learning gains, or educational opportunities. While TFA has internal metrics to evaluate corps member effectiveness, the external evidence evaluating TFA's impact on student academic performance is growing, but still limited.

The initial goal: Proximate student achievement. The notion that TFA should improve student achievement is the central component of the logic model, and gets at the most basic type of evaluation of TFA. The core of this intervention is focused on promoting student learning in a measurable way. Thus, as noted above, while it is important to consider how TFA might alter education policy, teacher education, and district and school operations more broadly, TFA should first be evaluated on the extent to which it increases student learning beyond what it would have been in the absence of the program. In other words, if TFA works as intended, we should see evidence of positive, proximate student outcomes among students in TFA classrooms.

Perhaps not surprisingly, this direct connection from TFA to short-term student outcomes is where the bulk of the research aimed at evaluating the effectiveness of TFA is situated. Several studies compare TFA teachers to either novice or experienced teachers in similar schools by examining end of year test scores. This evidence on program effectiveness is decidedly mixed. However the research designs used vary in quality and leave many unanswered questions even within this most heavily scrutinized portion of the logic model.

Researchers and district administrators have examined the relationship between TFA and short-term student outcomes with a variety of experimental, quasi-experimental, and descriptive research designs, in a variety of different contexts, and over a wide range of years. To date, only one study used an experimental research design that can examine the causal effect of TFA on student achievement. Glazerman, Mayer, & Decker (2006) evaluate the effect of TFA on student achievement in grades 1-5 using a random assignment design at six TFA sites, finding that students assigned to TFA teachers outperform students in non-TFA classrooms in mathematics, but not reading. The gains in math are equivalent to an additional month of math instruction across the included grade range, as calculated by the original authors using the ITBS norming conventions.[2] The effects are larger when TFA corps members are compared with novice teachers than veteran teachers and hold for comparisons across all certification types, with the exception of when first year TFA corps members are compared with all control teachers, where they are not statistically different. Importantly, this study compares TFA teachers to all types of non-TFA teachers in schools in which TFA operates, providing an appropriate counterfactual of the types of students and teaching these students would have received in the absence of TFA, including a mix of novice and experienced teachers. However, this research does not examine the effects of TFA teachers over time or in other grades. This work is currently being extended with an additional U.S. Department of Education i3 grant which supports the replication of these analyses with a second random assignment study of TFA that began in the 2012/13 school year (Teach For America, 2012).

A number of other studies use some form of quasi-experimental design to examine the relationship between TFA and student outcomes. Xu, Hannaway, & Taylor (2011) examine the relationship between TFA and student achievement in North Carolina high schools. They use a

student fixed effects design to compare performance across high school subjects in the same school year, exploiting the fact that students may have TFA teachers in some subjects but not others. Xu and colleagues find that TFA teachers are better than comparison teachers in science (ES = 0.19), with inconsistently positive results in math, regardless of the certification and experience level of the comparison teachers. However, their analyses are restricted exclusively to high school.

Other quasi-experimental studies rely on some combination of school, year, and grade fixed effects. For example, Boyd, Grossman, Lankford, Loeb, & Wyckoff (2006) compare the effectiveness of novice teachers of grades 4-8 who enter through a variety of different pathways. Using data from New York City, they find that teachers who enter through alternative pathways with reduced coursework, including TFA, outperform teachers who complete university preparation programs in both math and reading, but that most of these differences fade out as the cohort of teachers gains experience.

In addition, several states and school districts have produced their own reports, sometimes using a quasi-experimental design, and sometimes using linear regression with adjustments for observed student and teacher characteristics. Here, the comparisons are made relative to a variety of non-TFA teachers, some in the same schools and districts as TFA teachers and others including non-TFA schools and districts where students had no opportunity to be placed in a TFA classroom, and the conclusions are decidedly mixed. Raymond et al. (2001) find that students in TFA classrooms in grades 4-8 outperformed those in non-TFA classrooms in Houston, particularly in math in elementary and middle school. Darling-Hammond, Brewer, Gatlin, & Vasquez Heilig (2005) reanalyzed these data, finding contradictory results which suggest that TFA teachers are less effective than certified, experienced teachers, and slightly

more effective than other novice teachers, however these earlier studies compared TFA teachers to all teachers within the district and did not use school fixed effects to make comparisons within the same schools, which suggests that they were not using the most appropriate counterfactual teachers in their comparisons. More recently, Ware et al.'s (2011) evaluation of TFA effectiveness in grades 3-12 in Texas once again suggests a positive relationship between having a TFA teacher and student achievement, particularly in math, and especially for disadvantaged, minority students. Like Darling-Hammond et al. (2005), Laczko-Kerr and Berliner (2002) find that TFA teachers in grades 2-8 in Arizona are less effective than certified teachers and no different than other under-certified teachers. However, these estimates are generated by comparing student achievement across several districts, some of which did not have any TFA teachers, and identifying comparison groups using a test for district sameness, which suggests dubious counterfactual comparisons at best. By contrast, an evaluation of TFA in grades 3-12 Charlotte-Mecklenburg in the 2008 and 2009 school years finds that TFA teachers are more effective than other novice teachers, but not more experienced teachers (Schoeneberger et al., 2009). Similarly, in grades 4-9 in Louisiana, Noell & Gansle (2009) find that TFA teachers outperform traditionally certified novice teachers in math, science, English language arts, and reading, but not social studies, and their students' scores are not statistically different from those of veteran certified teachers. Finally in grades 3-12 in North Carolina, TFA teachers were found to outperform University of North Carolina (UNC) trained teachers on five of nine comparisons, particularly in high school and middle school, and sometimes by a wide margin (Henry et al., 2010), however, this study only compared TFA teachers relative to those trained by UNC preparation programs, and not relative to all counterfactual teachers in TFA schools.

Several conclusions can be drawn from the research examining the direct link between having a TFA teacher and proximate student outcomes. While the quality of the research designs varies greatly, taken at face value, TFA appears to be working well in some subjects and not in others. In addition, TFA teachers appear to have a more positive impact on math and science, and less so on reading or language arts. Finally, in most cases, TFA teachers outperform other certified and under-certified novice teachers, and they occasionally outperform veteran teachers as well. This suggests that TFA may not have a uniform effect across different sites, across different subjects, and across different grade levels. In other words, it is important to examine whether TFA has heterogeneous impacts along these dimensions. In addition such comparisons should compare TFA teachers relative to the appropriate counterfactual teacher in the same schools where TFA teachers are assigned, and make multiple comparisons across comparison teachers of different certification levels. Studies 1-3 of this dissertation examine different types of heterogeneity in this relationship.

The new goal: Altering long-term educational trajectories. More recently, TFA has extended its goals from producing short term learning gains to enacting longer-term change in students' academic trajectories (Kopp, 2011). While TFA has long discussed its desire to have longer-term impacts, now it is formally a part of how TFA judges the success of its corps members. While this is an obvious extension of the short-term outcomes literature, no evaluations of TFA impacts on longer-term outcomes exist. The only evidence we have of any long-term impact is the presence of "second generation" corps members who were students of TFA corps members (Guter, 2012). However, we have no way of knowing whether or not these students would have pursued the same path without the presence of a TFA teacher, or how widespread a phenomenon this is. More importantly, we have no small- or large-scale accounting

of the longer term outcomes and experiences of students taught by TFA teachers. In future work, I will test whether having a TFA teacher predicts long-term student outcomes that are measurable using school district data.

Classroom practices. While many studies utilize a basic framework for TFA that examines test score differences between students who did and did not have a TFA teacher, the underlying logic behind TFA is more elaborate. At its root, the idea behind TFA is that something happens in TFA classrooms that impacts student achievement. These teacher behaviors may consist of many things, for example instructional design, content knowledge, behavior management, interactions with parents, or seeking out community resources. The logic is that these teacher behaviors in TFA classrooms are somehow distinct from those in non-TFA classrooms and that these behaviors have differential impacts on student outcomes. Some research and work from the popular press examines teacher behaviors in TFA classrooms and describes some differences in TFA classroom planning and teaching, but it is very limited and not necessarily representative of most TFA behaviors (Foote, 2009). For example, Schoeneberger et al. (2009) observe 32 non-randomly selected TFA and non-TFA classrooms and identify several differences in terms of teacher instruction and classroom management. TFA classrooms have higher levels of respect in the teacher-student dynamics and dynamics between students, a larger variety and consistency in the use of classroom management strategies, a greater efficiency in classroom procedures, and a greater use of open-ended questions and probing for comprehension, a greater emphasis on real world connections. As they do not specifically examine achievement differences across the observed classrooms, it is not possible to link these practices directly to student outcomes, however, this is a first step for the type of

research that will help to identify the mechanisms at work in TFA classrooms that do and do not promote student learning.

Unpacking the TFA Intervention

The process by which TFA identifies and trains corps members is also complicated and under- investigated. The process, product, and organization of TFA are an amalgamation of several distinct components, all of which could have independent, and potentially even countervailing, effects. These are: recruitment, training (which is itself comprised of the pre-service 5 week institute and the credential courses), coaching from TFA staff, professional development once corps members are in their teaching placements, and volunteer peer mentoring from more senior corps members and alumni. While TFA regards these components as essential to acquiring and training good teachers, little is known about how these processes work independently, or in combination, to affect TFA teacher practices or student learning. While it is likely that each part contributes in some way, the precise role of each component in shaping teacher practices and producing student outcomes needs additional scrutiny, as this might enable other alternative and traditional certification programs to emulate some of the components associated with positive effects observed on student achievement.

Selection. Selection is an important aspect of TFA, particularly because of its connections to the larger arguments about teacher effectiveness. TFA chooses corps members from a highly selective pool of applicants, who are primarily recent college graduates. Many of these applicants come from elite colleges including Ivy League schools, top state universities, and historically black colleges and universities (Dobbie, 2011). The applicants, particularly those who are selected, differ from the majority of teachers who enter teaching through standard certification processes and many other alternate certification pathways as well. Decker, Mayer,

& Glazerman (2004a) detail these differences, which include the competitiveness of the college attended, SAT scores, and number of college level math courses taken, and indicate that TFA attracts an elite group of individuals into its program. Accounts by Foote (2009) describe the process, highlighting how TFA identified several applicants it ultimately accepted. These applicants had high ratings on such qualities as influencing and motivating others, perseverance, respect for others, and achievement, which are part of seven TFA selection criteria. In addition, work by Dobbie (2011) finds that higher rankings on some of the selection criteria, including leadership experience, perseverance, academic achievement, and commitment to the TFA mission, positively predict student achievement.

Training: Institute. Applicants who are extended an offer and accept it become corps members and begin their two year commitment with training. Corps members receive two primary types of teacher training. Initially, they enter an intensive five-week summer training program, referred to as "Institute". During this time, corps members attend courses about pedagogy, content, classroom management, diversity issues, and many other topics. In addition they work in a team of four to teach summer school for four weeks. This training is substantially less than what teachers receive if they attend a standard credential program before their first teaching position, but it is more than what many novice teachers who enter through other alternative pathways receive. Foote (2009) provides an in depth observational account of Institute and Veltri (2010) describes Institute using retrospective teacher interviews. Both note the frenetic pace of Institute and the vast amount of information that corps members are expected to absorb and implement. They also document the subsequent struggles of new corps members when transitioning from Institute to their own classrooms in their first months of teaching. However, no research to date provides a detailed description of Institute training sessions or

linked aspects of Institute, such as corps member performance ratings, to student outcomes, thus it is difficult to assess whether it is effective in shaping practices that positively affect student achievement.

Training: Credential courses. Once corps members are in their placement sites, they also attend standard credential courses at an outside university. These courses are taken at night and on weekends and in some programs corps members can receive a master's degree is they complete additional coursework. Credential programs also often include classroom observations by faculty or mentor teachers from the credential program. This part of TFA training might not differ from what other non-TFA credential students in the same programs experience, although it could because sometimes TFA teachers receive TFA-specific courses from credential programs to accommodate their work schedule. Foote (2009) and Veltri (2010) describe some components of the credential training provided to TFA corps members from their observations and interviews. Foote (2009) provides some suggestion that TFA teachers do not find their credential courses very useful. This conclusion is supported by survey results from H. Carter, Amrein-Beardsley, & Hansen (2011), who find that TFA teachers generally offer more critical evaluations of their credential courses than non-TFA teachers. In contrast, Veltri (2010) details her role as a university instructor and teaching coach for TFA teachers, highlighting how it benefitted corps members. She suggests that her mentorship, which includes classroom observations, lesson modeling, and co-teaching in corps members' classrooms, is a particularly important aspect of support and development for TFA teachers. She also argues that TFA teachers do not have many sources for this type of sustained, involved mentorship, but that they need more, although this conclusion is based purely on Veltri's first-hand account and has not been corroborated by other sources. However, as with Institute, existing research does not make

a connection between coursework quality and corps member classroom behavior, changes in corps member teaching, or student outcomes.

Coaching. In addition to the pre-service Institute and the concurrent credential course work, TFA has three other avenues through which it tries to further develop and support its corps members. The primary method is through observation and coaching from a TFA staff person. This position was previously called a Program Director (PD) and has recently been reframed as a Manager of Teacher Leadership and Development (MTLD). MTLDs observe Corps Members several times a year, provide them feedback and coaching, help them to analyze student data and plan accordingly, and help them to set instructional goals for their classrooms (Sawchuk, 2009). While this process has been described by Foote (2009), it has not yet been examined systematically, nor has it been connected to teacher practice or student outcomes.[3]

Professional development. Corps members are required to attend several weekend professional development meetings with other corps members in their region. The instructors for these meetings are often TFA alumni, but also include outside instructors. In these sessions, corps members meet individually or in small groups with alumni or MTLDs for planning and coaching sessions. As with the other components of the TFA training and mentoring conceptual model, only a small amount of research examines these interactions. Veltri (2010) describes that some workshop facilitators lacked important content and pedagogical knowledge. She notes that several students did not get the support they wanted from these experiences, although it is unclear if this sentiment is held by all corps members. Regardless of quality, the intention of these coaching sessions is to enable corps members to talk about improving instruction to support student learning and performance. However, little is known about whether or not teacher practices or student outcomes change as a result of these experiences.

Mentoring. Finally, TFA also relies on alumni and second year corps members to provide mentoring to newer corps members. Corps members are often matched within schools, when possible, or within districts, to a TFA teacher mentor. Sometimes these mentorship activities are coordinated across a group of corps members in a regional Learning Team (Teach For America, 2013a). Foote (2009) describes many such mentoring interactions among high school teachers within the same school and subject area. These sessions often center on lesson planning and coordinating student support. While Foote describes this mentoring is beneficial for some, others engage in little of it, primarily because they find it of limited use and choose not to attend. No other research examines this peer mentoring, and thus is it not possible to know how common or extensive these experiences are.

Externalities

TFA may produce a variety of other impacts that fall outside of this conceptual model, and an accompanying body of research evaluates them. Most of this research focuses on the byproducts of TFA that occur over the course of the TFA program, rather than on the core elements of the TFA program, such as TFA teacher training and practice or student outcomes. These outcomes and processes are what I consider to be the "secondary" effects of TFA which focus on the unintended consequences of TFA and reach beyond the core parts of the TFA model, and include, for example, the impact of hiring TFA teachers on more general teacher recruitment in districts that staff TFA teachers or the impact of TFA on the behaviors and attitudes of the corps members once they have completed the TFA program. Although they may be important and may overshadow the primary intent of the program in some situations, they are not part of the central processes that define TFA.

Below, I briefly review the potential and observed positive and negative externalities of the TFA program. These are sometimes intended, sometimes unintended, positive and negative consequences of TFA, beyond the impacts of recruitment, training, and teaching on student outcomes. While these positive and negative externalities are outside the conceptual model and are not part of the key relationship between TFA and student outcomes, they are also important to consider when evaluating the larger effects of TFA. Even if the positive externalities are very large, if TFA does not have positive outcomes on student performance, then TFA should not be considered a success. In addition, if TFA is positively related to student outcomes, but contributes to sizable negative externalities, then this cost of the program must be weighed relative to the size of the benefits. These secondary consequences of TFA must be examined and considered when considering the full impact of TFA.

Positive externalities. Several of the potential positive externalities of the TFA program are intended by the program itself, but are beyond its primary focus. One observed positive externality that results from participation in TFA is a change to corps member beliefs. Dobbie and Fryer (2011) identify changes to corps member beliefs as a result of their teaching experiences. They find a discontinuity in which, relative to TFA applicants where were not accepted, TFA corps members have increased racial tolerance and beliefs in life chances of poor children. While these changes to corps members may impact student outcomes in a positive way, they do so outside of the direct path from TFA to classroom teaching to student outcomes.

Another potential positive externality of the TFA program is the long-term participation of corps members in the field of education. While some alumni stay on as classroom teachers, principals, and district administrators (Teach For America, 2013b), many alumni become involved in some other aspect of educational leadership. Dobbie and Fryer (2011) and Higgins et

al. (2011) examine the career paths of TFA alumni finding that alumni are more likely than those who did not enter TFA or dropped out to be working in the field of education. They also find that many TFA alumni are leaders in education policy, educational organizations, and educational entrepreneurship.

Lastly, one potential positive externality is the formation of similar programs in the US and globally, like many state and city Teaching Fellows programs and international TFA affiliates like Teach For India. Such programs have adapted the TFA model to fit their local circumstances. They have incorporated elements of the TFA training and support program into their own operations. However, this is only a positive externality if TFA and these other similar programs have positive effects on student outcomes. Some of the research that has examined TFA has also looked at the relative effectiveness of the alternative certification programs that emulate TFA. Their track record is mixed. Teachers certified through these pathways are sometimes just as effective as TFA teachers, and sometimes less effective than TFA teachers; in some cases they are more effective than university prepared teachers and in some cases no different (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2008; Kane et al., 2008)

Negative externalities. Critics and supporters of TFA also identify several potential negative externalities. One of the most often cited negative externalities of the program is its high cost as a solution to teacher shortages in high-need school districts. Several papers have noted the cost to recruit and train TFA teachers (Vasquez Heilig, Cole, & Springel, 2010; Veltri, 2010) and suggest that the same amount of money could be put to greater use in the hands of a school district or coordinated effort at improving university teacher preparation programs. While

they note the high costs, no one has conducted a comparative cost-benefit analysis examining the impact the same amount of money might have if spent differently.

Another noteworthy potential negative externality is TFA teacher turnover. TFA has no requirements about whether TFA teachers will leave the classroom immediately after completing the program or not, although they do explicitly promote alumni participation in the field education in their core model. However, because the program only requires a two year commitment from its corps members, it contributes to teacher turnover among novice teachers. In addition, TFA provides specific support to its alumni to apply for graduate school and to search for teaching and non-teaching jobs outside of their original assignment. Ethnographic work suggests that turnover is prevalent across many TFA sites (Foote, 2009; Veltri, 2010). This is supported by larger-scale quantitative work. Noell and Gansle (2009) find that TFA teachers had lower five year persistence rates than new teachers with standard teaching certificates. Kane et al. (2008) find that TFA has high rates of turnover relative to teachers trained through other pathways, although they suggest that the impact of turnover on achievement is modest, but Ronfeldt, Loeb, and Wyckoff (2013) find a more substantial negative impact of teacher turnover on student achievement. Donaldson and Johnson (2010, 2011) find that turnover is more likely among corps members who are placed into difficult conditions with a poor match between their academic background and the subject they are assigned to teach.

An additional potential negative externality is that TFA teachers end up displacing more experienced teachers. This issue was heightened by the financial crisis when many districts that take TFA corps members had to cope with budget shortages by mass rounds of layoffs. Veltri (2010) documents displacement in several TFA sites across several states. An evaluation of the prevalence and impact of teacher displacement has not yet been done. Given the mixed results

about the effectiveness of TFA teachers relative to those with higher levels of certification in the same schools, it is unclear what the impact would be on student outcomes. Because it is unclear which teachers are being displaced, the impact on students is ambiguous.

Finally, McAdam and Brandt (2009) examine the consequences of participation in TFA on civic attitudes and behaviors. While this has the potential to be a positive externality, it is instead an observed iatrogenic effect of TFA participation. While TFA participants outscore non-matriculants and program dropouts on attitudes about civic commitment, this is often related to their ongoing participation with TFA itself. In addition, TFA graduates lag behind non-matriculants in terms of service activity and both non-matriculants and dropouts in civic and political activity, and in voting.

Summarizing the Research & Conceptual Model

Teach For America is a rapidly expanding and increasingly influential alternative teacher recruitment and training program with an ambitious goal. It hopes to affect large scale change in the educational opportunities of low-income students across the country. This conceptual framework is intended to help researchers and policy makers develop a better understanding of TFA, and highlight our lack of knowledge about how key aspects of the program might affect student learning. On balance, the evidence suggests that TFA has positive effects on mathematics achievement, but little to no effect on reading, across a variety of ages. Developing a better understanding of the components of TFA that are and are not boosting student achievement will allow TFA to improve and other programs to learn from its successes while avoiding its failures. Below, I summarize the evidence, highlighting what TFA appears to be doing well, where it needs to improve, and where more evidence is needed before even tentative conclusions can be reached.

Strengths of the Program

Based on the existing research and viewed through the lens of the conceptual model, TFA has some positive impacts. Most notably, TFA teachers are generally found to be more effective in math and science than traditionally and alternatively certified novice teachers, and even some veteran teachers. While not observed in every evaluation, this finding has been replicated across a variety of locations and a wide range of grade levels using rigorous research designs. Although the impacts are small, TFA has a positive impact on short term student achievement in some subject areas.

In addition, the program has positive externalities on the beliefs and careers of its corps members. Participating in and completing TFA increases the likelihood of participation in education-related careers. TFA participation has a positive impact on the racial attitudes of its corps members, their beliefs about the future possibilities for their students. These positive externalities are certainly much less important than the direct impacts the program has on students, but they are positive nonetheless.

Finally, TFA has helped elevate issues related to educational inequality to the forefront of national policy conversations. While people may have mixed feelings about its model for recruiting and training teachers, its business practices, and its political influence, it has no doubt played a role in focusing this conversation on issues of teacher quality and access. It has elicited increased interest in examining and addressing issues related to teacher shortages, teacher training, and the inequity of opportunities that stem from socio-economic and racial residential segregation and inequality.

Weaknesses

While TFA teachers are more effective in math and science, they do not appear to be producing large enough gains for students in TFA classrooms to catch up with the average US student. That is, while TFA teachers helped their students learn more than non-TFA teachers in similar classrooms, these gains did not erase socio-economic achievement gaps. Likewise, the lack of differences between students in TFA and non-TFA classrooms in reading and language arts across nearly every study is troubling. While it is good to know that TFA does not worsen reading performance, given the amount of money invested in the program[4], it is also disappointing that it has no impact on reading.

Another clear issue is the high teacher turnover among TFA alumni. TFA placement schools already have higher turnover than most schools and TFA contributes to the perpetuation of this phenomenon. While it is unclear what the cost is to student academic development and performance, and indeed some suggest its impact on achievement is modest (Kane et al., 2008), it is disruptive to school communities. Thus, this cost should be weighed when considering the tradeoffs of TFA.

Teacher displacement is another potential issue, however, the evidence here is less clear. If TFA is displacing teachers with more seniority who are more effective than TFA teachers, then this is certainly a cost of the program. However, seniority is not a guarantee for quality and effectiveness, so it is difficult to determine the implications of displacement for students without additional evidence.

One final, potentially surprising, externality is the decline in civic engagement among TFA alumni. It is not clear what the larger implications of this finding are, but it certainly dampens any potential benefits from improvements in racial attitudes.

The Unknown

In examining the research on TFA, the conceptual model highlights how many studies have examined the relationship between TFA and proximate student outcomes. This work is full of mixed and contradictory results, suggesting the need for additional replication with strong experimental and quasi-experimental research designs, paired with classroom observations to uncover mechanisms behind successes and failures. These contradictory findings make it difficult to be certain about program impacts and to pinpoint which aspects of the program contribute to its success or failure.

One particularly noteworthy aspect of this information deficit is that most work examining the impact of TFA on student outcomes does not consider the possibility that the impact of TFA may be heterogeneous. In other words, even if TFA is beneficial on average, it may actually have positive effects for some and negative effects for others which offset one another. Outside of a handful of studies that compare differential impacts by student subgroups (Glazerman et al., 2006; Xu et al., 2011), nothing is known about this. This is a potentially promising avenue for future research that may help to reconcile disparate findings.

In addition, extending this work to analyze long-term student outcomes is important. While TFA teachers may positively impact student growth in state test scores, what TFA and we as researchers and educators care more about is whether these gains transfer to longer-term achievement and academic outcomes, such as high school graduation and GPA. An analysis that expands beyond test scores and looks at longer term outcomes is an important next step for deepening our understanding of the effects of TFA on students.

While the connection between having a TFA teacher and short-term student outcomes is the most researched area, other parts of this conceptual model yield interesting insights and point to promising directions for developing a deeper understanding of how and whether this

organization benefits student learning. The model highlights that we know relatively little, particularly regarding the inner workings of TFA training and TFA classroom teaching. While only a small portion of the research connects specific aspects of the TFA model with student outcomes, understanding which aspects of the TFA program are important could help to refine TFA and other similar teaching programs.

Currently, selection is the only individual component of the TFA intervention that has been specifically linked to student outcomes. Dobbie (2011) suggests that selection is an important and predictive aspect of student performance. However, although this work suggests that the TFA approach to selecting teachers can contribute to student success, the precise mechanisms at work are unclear. This suggests that future explorations of the relationship between TFA and student outcomes should attend to selection criteria.

In addition to selection, mentorship of corps members, both by credential staff and faculty, and by TFA staff is an area that needs additional exploration. While selection is an important part of the TFA model, TFA recognizes the need to mentor corps members as they are becoming teachers, particularly in such a rapid fashion. Despite this fruitful area for research, which might provide insights about the mentoring of teachers more broadly, very little work has examined how mentoring works in TFA. A few notable exceptions underscore the need for more work in this area. Veltri's (2010) corps member interviews and coaching sessions suggest that, while the Institute is intended to support corps members' skill development, the coaching that occurs after Institute might be more consequential for refinement of teacher practice and daily classroom interactions. While Veltri does not observe the coaching interactions between corps members and TFA staff (which she suggests are not frequent or supportive enough to promote development of pedagogical skill), she does describe the types of mentorship she herself

provided. She argues that for many of the teachers in her study, her coaching was a key to developing their own management and pedagogy. This suggests that observations and comparisons of the various types of mentoring and coaching post-Institute would be useful, particularly if they can be connected to student outcomes.

Likewise, to deepen our knowledge base about the coaching TFA teachers receive, it is also important to investigate the teaching practices utilized by TFA teachers. Ideally this research will compare TFA teaching practices to those of non-TFA teachers in the same schools. The Charlotte-Mecklenburg report by Schoeneberger et al. (2009) identifies some observable differences in instruction between TFA and non-TFA classrooms, and that the TFA teachers exhibited more behaviors that might support student learning, including good classroom management practices and using supplemental materials and new ways of teaching to keep students engaged. Similar studies, with larger, representative, and possibly randomly selected samples that are linked to student outcomes would clearly be of great value.

Theorizing Variation in the Effects of Teach for America

Having reviewed the literature on TFA above, I next provide an overview of research that provides us with insight into why we might expect the effects of TFA to vary across different students. I I draw on research from a wide variety of topics, mostly examining non-TFA interventions and processes, that can help us understand why the effects TFA might not be uniform. This previous research provides a myriad of countervailing paradigms, with some research suggesting that TFA might have larger effects in elementary school, others providing reasons why we might expect larger effects in middle school, and still others in high school. Variation in the effectiveness of TFA teachers could result from differences in the importance of content knowledge, the number of contact hours, the developmental timing of having a TFA teacher, and the effectiveness of the TFA training for a particular age group and school type. It is

also possible that the quality of TFA teachers varies little, and that variation in the effects of having a TFA teacher is attributable to differences in the quality of the counterfactual teacher. Counterfactual teacher quality may vary because of differences in academic background and teaching skill, match between certification area and teaching assignment, the quality of certification and training programs, and rates of turnover and the desirability of particular teaching assignments. While it is not possible to adjudicate between all of these potential mechanisms behind the variation in the effectiveness of TFA, particularly as it is likely that many could be operating simultaneously, these mechanisms can be used to theorize about the levels and subject areas in which TFA is likely to be more or less effective and why.

Content knowledge. Differences in the importance of highly developed content knowledge might be one axis along which the relationship between TFA and student performance can vary. Several studies have identified a positive relationship between teacher subject-matter knowledge and competence, as measured by teacher certification exams, and student achievement (e.g., Boardman, Davis, & Sanday, 1977; Ferguson, 1991; Greenwald, Hedges, & Laine, 1996; Hanushek, 1972, 1986; Harbison & Hanushek, 1992; Mullens, Murnane, & Willett, 1996; Rowan, Chiang, & Miller, 1997; Strauss & Sawyer, 1986; Wayne & Youngs, 2003). In North Carolina in particular, this positive relationship between subject matter measured by teacher certification exams and student performance has been found to be positive and significant for math and English, but not science (Clotfelter, Ladd, & Vigdor, 2010). Others link verbal or mathematical ability to gains in student achievement (Ehrenberg, Goldhaber, & Brewer, 1995; Ferguson & Ladd, 1996; Hanushek, 1971; Harbison & Hanushek, 1992; Mullens et al., 1996; Rowan et al., 1997). Still others argue that accomplished teachers need to master content knowledge, pedagogical content knowledge, and curriculum knowledge to effectively

teach students (Shulman, 1986, 1987; Wilson, Shulman, & Richert, 1987), and research examining the relationship between content knowledge for teaching and student achievement suggests that underlying content knowledge positively impacts student achievement gains in the area of math (Hill, Rowan, & Ball, 2005). Thus, differences in teachers' subject matter knowledge, verbal or mathematical skills, or content knowledge, pedagogical content knowledge, and curriculum knowledge might all positively impact student achievement.

Prior research demonstrates that, on average, TFA teachers come from more selective educational backgrounds and possess higher levels of advanced content knowledge than teachers certified through other pathways. TFA teachers are more likely to come from more elite universities and have taken more rigorous undergraduate coursework than non-TFA teachers (Boyd et al., 2006; Decker et al., 2004a), and outperform teachers trained through other pathways on certification exams testing general knowledge, and liberal arts and science concepts (Boyd et al., 2006). Specific comparisons of math qualifications and outcomes of teachers trained through TFA versus other alternative and traditional pathways suggest that TFA teachers have stronger math credentials and achievement, including test scores, coursework, and competitiveness of degree granting institution (Boyd et al., 2010). In high school, TFA teachers also have higher scores on their general content and math-specific certification tests (Xu et al., 2011).

Thus, given TFA teachers' higher levels of content knowledge, we would expect students who have TFA teachers in related course areas to outperform students in non-TFA classrooms. However, the degree to which this advanced content knowledge translates to student achievement might vary across subject and school level if specific content knowledge is more or less central to instruction across different levels and subjects. As a result, we would expect to see

a positive relationship between having a TFA teacher and student achievement in math at all grades, but the magnitude would be largest in high school, moderate in middle school, and smallest in elementary school. A similar pattern would likely emerge in science as the more advanced content knowledge also has the highest payoff as the course content becomes more complex.

In language arts, the pattern might be somewhat different. While TFA teachers likely have more experience with analytical writing and literature, it is less clear that this supports the type of language arts content that is being taught at all levels. While we might expect experience with literature content to have a positive impact on students in high school English classes, it is less clear if this translates into a teacher skill set that supports training for remedial writers and readers. It is also unclear whether any specific part of a TFA teacher's background includes training for instructing early readers and struggling readers in elementary school, particularly since few TFA teachers major in education in college (Decker et al., 2004a). In language arts, TFA teachers may actually be at a disadvantage relative to non-TFA teachers in elementary school language arts because they do not have adequate training to teach phonics and intervene on behalf of students who do not have strong literacy foundations. Thus, in language arts, assuming some content knowledge about the mechanics of reading is important, it seems possible that TFA teachers might not have a positive effect at any school level, but if they do, it would be expected to be largest in high school, moderate in middle school, and non-existent in elementary school.

While TFA teachers' content knowledge and overall verbal and mathematical abilities might be greater than those of non-TFA teachers, their limited training experiences might put them at a disadvantage in terms of pedagogical content and curricular knowledge. In subject

areas and grade levels where either pedagogical or curriculum knowledge is especially important, such as with early reading instruction, TFA teachers may not be as effective as their more experienced counterparts.

School context and organization. An alternative mechanism that may lead to variation in the effectiveness of TFA teachers across school levels is differences in school context and organization. Differences in school structure across school levels create environments in which teachers can be more and less supportive of students. Several elements of school structure might contribute to the ability of teachers to effectively support student engagement and achievement, including the duration of the class time a student spends with a particular teacher, whether the student is in self-contained classrooms or has a variety of teachers each day for specific subjects, and the size of the school population.

One way that teachers' impact on achievement might vary is through the amount of time students spend with a particular teacher. Results from interventions aimed at extending the school year and the school day suggest that increased school time has positive impacts on student achievement (Bellei, 2009; Skandera, 2012). In addition, numerous academic interventions ranging from early childhood education to remedial adult education, to school access suggest that increasing the duration of exposure to these interventions positively impacts academic test scores, language development, and other school-readiness skills (Armecin et al., 2006; Banerjee, Cole, Duflo, & Linden, 2007; Behrman, Cheng, & Todd, 2004; Chin, 2005; Duflo, 2004). Finally, variation in time allotted to instruction in a given domain, such as math, also has an impact on domain-specific achievement in the elementary years (Brown & Saks, 1986; Entwisle & Alexander, 1999). Thus, to the degree that TFA teachers have a positive impact, the greater

the amount of exposure students have to a TFA teacher, the greater we might expect that impact to be.

One way that this increased exposure might operate is if self-contained classrooms allow effective teachers to integrate the curriculum across the school day, reinforcing concepts across subject areas and providing ongoing emotional and academic support. Developmental science suggests that children learn best in settings in which curriculum is integrated and subject-matter is interrelated (Bredekamp & Rosegrant, 1992; Cobb, 1994; Dewey, 2004; Piaget & Inhelder, 1969). Thus, to the degree that self-contained classrooms not only provide more exposure to TFA teachers, but also allow teachers to integrate their curricula, we might expect that skilled teachers in these settings might be particularly beneficial.

Eccles and Midgley (1989) have also identified that strong teacher-student relationships, which are easier to build in self-contained classroom settings, are positively related to student achievement. Eccles and Midgley also underscore the importance of stage-environment fit, noting that self-contained classrooms, in which a single teacher is responsible for instruction across subjects, better meet the emotional, behavioral, and academic needs of students in elementary school and would do so for middle school students as well. Self-contained classroom is a better developmental fit for children and adolescents, which supports academic self-concept, motivation, and school persistence (Eccles & Midgley, 1989; Eccles et al., 1993; Eccles, Lord, & Roeser, 1996). Traditional middle schools are a particularly bad fit for adolescents because they are characterized by weaker teacher-student relationships which provide less autonomy for students and lead teachers feeling less effective instructionally (Brophy & Evertson, 1978; Eccles & Midgley, 1989; Midgley, Feldlaufer, & Eccles, 1989), and the transition to this organizational structure is particularly ill-timed with the developmental transition to adolescence

(Eccles et al., 1993). Alspaugh and Harting (1995) find that the transition from self-contained to multiple subject-specific classrooms has a negative impact on student achievement.

Students in school structures that are well-aligned with their developmental needs, such as self-contained elementary school and K-8 schools, might benefit most from having a TFA teacher, in part because the TFA teacher can take advantage of the supportive environment to focus more energy on developing high quality instruction and less on addressing organizational and behavioral challenges. This supposes that more effective teachers are especially beneficial in contexts where schools structures work in concert to support student achievement and allow teachers to tailor instruction throughout the school day to meet students' needs. Prominent examples of this type of environment are the KIPP schools, which provide a highly structured environment with increased exposure to high-quality teachers. Teachers in KIPP schools can devote substantial amounts of time to supporting student learning, allowing effective teachers to have substantial impacts on student outcomes (Angrist, Dynarski, Kane, Pathak, & Walters, 2010).

If TFA teachers are better able to take advantage of developmentally aligned school environments, then their impact is likely to be largest in the settings in which they have more contact with students and are able to be the most supported by the school environment. This is most likely to occur in elementary school, then in self-contained middle schools, and less likely in traditional middle schools and high schools.

Fit within TFA training and support. In addition, alignment differences in the TFA training and support model and teacher needs might impact the effectiveness of TFA teachers. Work examining the impact of alignment between professional development activities, instructional practices, and curricular reform suggests that instructional practices that are more

aligned with curriculum are positively related to student achievement (McCaffrey et al., 2001). If the TFA model is better adapted to the curriculum of a particular school level (e.g. elementary as opposed to high school), then this may impact the effectiveness of TFA teachers in several different ways. For example, TFA teachers at some levels may have summer school teaching experience in the grade and subject level that they teach during their two year assignment, while others may not have an exact match. If the training location where they teach summer school has different grade level needs than their placement region, then their training may not prepare them as well as teachers who are assigned a close or exact match. For example, if a summer school training district only needs summer school teachers for grades 1-5, but the TFA corps member is eventually assigned a self-contained 6th grade or kindergarten classroom, then the training might be less aligned than someone who teaches second grade in summer school and at their placement site.[5] A similar issue could occur for mathematics teachers in high school. Perhaps summer school needed only remedial sessions of algebra and geometry, but the TFA teacher is eventually assigned to teach calculus. This could also lead them to be less prepared than someone who teaches algebra in both summer school and in their placement.

TFA training may also contribute to differences in effectiveness once corps members are in their teaching placements based on the availability of mentor teachers and staff with experience in the same grade level and content area. TFA tries to place corps members in schools with other TFA corps members or alumni, however this is not possible in all cases. New TFA corps members are also put in contact with an alumni or second-year mentor in similar content areas if no match is available at the same school site. However, this match is still not perfect in many cases. A third grade teacher might be mentored by a fifth grade alumnus, and a high school biology teacher might be mentored by an alumnus who teaches chemistry. In addition to poor

matches with mentors, if the local staff tasked with supporting a particular TFA teacher does not include someone with experience in a similar grade level or content area, the level of mentoring might be worse than if there was a grade level and content match. For example, if a TFA teacher is placed in a bilingual classroom, but their TFA supervisor does not speak Spanish, they might not receive the same quality of supervision as other corps members in non-bilingual classrooms. While issues related to the fit of TFA training and support do not lead to predictions of TFA teachers always being more or less effective in specific grade levels and content areas, they do suggest that TFA teachers placed into new grade levels and subject areas that were not present in the region previously (or for several years) might less effective than those placed into subjects and grade levels with a strong presence in the region.

Developmental timing. The impacts of TFA teachers may vary depending on the developmental timing during which a student has a TFA teacher. Following Cunha, Heckman, Lochner, and Masterov's (2006) theories of self-productivity and complementarity, skill attainment and investment at earlier stages facilitate skills acquisition and productivity at later stages. While later investment is needed to maintain earlier gains, the returns to investing early are high.

Research examining the returns to many educational interventions compares the impact of programs targeting early childhood. A summary of such interventions by Currie (2001) suggests that some, such as Head Start, are better with a follow up in early elementary school. Others, like the Carolina Abecedarian Project, are more effective as an early childhood only program or when an early childhood intervention is paired with a primary school intervention. In contrast, receiving only the elementary school intervention was less effective (Campbell & Ramey, 1994).

While TFA has begun placing teachers in pre-K classrooms, this is only a recent development and comparisons across K-12 interventions are more relevant. However, research comparing interventions across school levels also suggests differential effectiveness that varies by developmental stage. For example, a meta-analysis of experiments evaluating the effects of career education interventions found larger effect of such programs targeted at elementary, rather than middle or high school students (Evans & Burck, 1992).

Thus, to the extent that TFA has a positive effect on student achievement, research examining the returns to early intervention would suggest that larger benefits will be found at earlier ages. In addition, TFA places a large emphasis on building relationships with students and their families. This type of connection might matter more at some ages than others, or may be missing in counterfactual classrooms more at some ages than others, which might contribute to differences across developmental stages.

Counterfactual teacher quality. Finally, differences in the effects of TFA teachers across developmental stages may be less a function of differences in TFA teachers, but instead may result from differences in the quality of counterfactual teachers at each stage. Overall, research suggests that 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). Evidence suggests that, while average teacher quality 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). This suggests that new teachers from higher-caliber backgrounds are becoming increasingly rare. Further, as TFA seeks to staff hard to fill positions in underserved districts, these schools may have particularly weak counterfactual teachers. This is because staffing is

typically done by seniority, and as teachers gain experience, they become increasingly likely to leave low-performing, high-minority schools in lower-income areas (Carroll, Reichardt, Guarino, & Mejia, 2000; Cha & Cohen-Vogel, 2011; Cohen-Vogel, Osborne-Lampkin, & Houck, 2013; Cohen-Vogel & Osborne-Lampkin, 2007; Darling-Hammond, 2003; Esch et al., 2005; Hanushek, Kain, & Rivkin, 2004a; Scafidi, Sjoquist, & Stinebrickner, 2006), so that non-TFA teachers in TFA placement schools are likely less qualified and less experienced than non-TFA teachers in more affluent, higher-achieving schools. Evidence from North Carolina confirms that high-poverty schools, including those where TFA teachers are placed, have more inexperienced teachers from less competitive undergraduate institutions, with below-mean licensure test scores and fewer teachers with National Board Certification (Clotfelter, Ladd, & Vigdor, 2007a; Clotfelter, Ladd, Vigdor, & Wheeler, 2006).

However, while non-TFA teachers in TFA placement schools come from weaker backgrounds than TFA teachers and non-TFA teachers in more affluent schools on average, the quality of these counterfactual teachers may vary considerably, which could be the result of several different factors including: teacher skills, fit between teacher certification area and placement subject, quality of teacher training, and turnover rates.

One potential source of variation in teacher quality across school level and subject is differences in the academic background of non-TFA teachers, both of which should contribute to their skills as a teacher. It may be the case that non-TFA teachers at particular stages have more rigorous educational backgrounds and experience that make them stronger teachers than the non-TFA teachers in other developmental stages. For example, teachers with higher level math skills typically teach high school math (Boyd et al., 2012), and most high school math teachers have several years of college-level math coursework and many majored in math or a quantitative

subject (Ingersoll, 1999). However, as has been documented in previous research, few elementary school teachers take advanced math or courses in college (Book & Freeman, 1986), and few of them major in math, science or any quantitative subject (Ingersoll & May, 2012). Thus, in high school math, the non-TFA teachers' instruction might compare favorably relative to that of TFA teachers, while in elementary school, TFA teachers' math instruction might be substantially better than that of non-TFA teachers. While the TFA teachers themselves might have similar levels of math experience regardless of school level, the counterfactual teachers differ dramatically in their relative level of experience and comfort with math across school levels. This may lead TFA teachers to appear more effective in elementary math than high school math, simply by virtue of the relative facility with math of the counterfactual teacher.

Another way that teachers might differ in quality across school levels is through the match between their teaching assignment and their degree of study in their undergraduate or master's program or their subject area of certification. Work summarized by Ingersoll (1998) suggests that many teachers are not teaching in fields that correspond to their undergraduate degree. Even in advanced fields, such as secondary math and science, as many as a third of teachers were instructing courses with out-of-field degrees (Ingersoll, 1999). Thus, to the extent that counterfactual teachers' degree-assignment match is worse at certain developmental stages, this may indicate that they have better training and lead them to perform comparatively worse than TFA teachers. In North Carolina, one way this fit is measured is through subject-specific certification, which can be earned through passing a subject specific certification test. Clotfelter, Ladd, and Vigdor (2010) find a positive relationship between teacher certification in high school math and English and students' scores in those subjects, but not others. This suggests that, in

addition to variation across developmental stages, this relationship might vary across content areas.

Alternatively, it may not simply be the match between teacher degree and teaching assignment. It may instead by the case that teacher training programs are more effective at training teachers for certain subjects and grade levels. That is, if teacher training programs have a strong conception of what an effective high school math teacher needs to know, but little sense of what an elementary language arts teacher needs to master, then we would expect high school math training to produce more successful teachers than elementary language arts training. Research in North Carolina finds some evidence of this. For example, Clotfelter, Ladd, & Vigdor (2010) find that receiving a master's degree has a small positive relationship with student achievement in high school, while it has a negative relationship in elementary school. Similarly, if state certification boards know what teachers need to know in order to be successful teachers, then state certification exams should also be more predictive of student outcomes. Evidence suggests that training, certification, and student learning are more closely aligned in some subjects than others. In North Carolina, high school teachers' certification test scores in math and English are positively associated with their students' scores on end-of-course tests in the same subjects, and which is larger than the positive relationship found between teacher certification tests and student achievement in elementary school (Clotfelter et al., 2007a, 2007a, 2010). If additional sources of training and certification, such as National Board Certification, are better calibrated to what teachers need to know to be successful for some subject areas and school levels than others, then this can also create variation in the quality of the counterfactual teacher. North Carolina is a national leader in promoting and providing incentives to teachers to pursue National Board for Professional Teaching Standards (NBPTS). It offers a 12 percent pay raise to

teachers who obtain certification. Clotfelter, Ladd, and Vigdor (2010) find that high school teachers' teaching improves as a result of undertaking the Board Certification process whereas no such effect is found for elementary school teachers who obtain Board Certification. While few TFA teachers obtain National Board Certification, the varying returns to this certification process suggest that the counterfactual high school teachers may become comparatively stronger teachers after having gone through the process.

Finally, teacher quality may vary across developmental stages because of differences in the rate of teacher turnover across different school levels. Previous research suggests that teacher turnover is greater among teachers with higher test scores and that high school chemistry and physics teachers are more likely to leave than teachers in other content areas (Murnane & Olsen, 1989, 1990). A meta-analysis of teacher turnover confirms these findings. Borman and Dowling (2008) find that high school math and science teachers have higher rates of turnover than elementary school teachers, who in turn have higher rates of turnover than non-math and science high school teachers. Little empirical evidence exists about the level of turnover and difficulty of staffing middle schools, however anecdotal evidence suggests that middle school grades are the most difficult to staff and are plagued with teacher turnover (Useem & Neild, 2001), and a survey of education majors from North Carolina and Virginia suggests that middle schools are the least desirable due to concerns around discipline and adolescents' attitude problems (M. S. Carter & Carter, 2000). Given that TFA typically places teachers in hard-to-staff schools, where veteran teachers are more likely to transfer to other schools and teachers of all experience levels are more likely to quit, these schools are also more likely to have inexperienced teachers in general, but particularly in hard to staff subjects. This suggests that non-TFA teachers in math and science in high school, as well as middle school teachers, are more likely to be novices than

elementary school teachers and high school English teachers. If experience is an important factor, this could lead to a larger TFA advantage in high school math and science and middle school.

Synthesis and Overview of Literature Review

This literature review presents a conceptual model to synthesize and analyze the existing literature on the effects of TFA. It uses an understanding of how TFA is supposed to operate and then incorporates other secondary effects of TFA to inform our understanding of the variety of impacts TFA has on student outcomes, TFA corps members, and the field of education more broadly. Examining the corpus of research in this way underscores the presence of contradictory findings and differences of opinion about the merits of TFA. It also highlights that many areas for researchers to investigate more fully, which could potentially have major implications for the future of education reform.

In addition to this most basic understanding of TFA, this summary of the evidence identifies positive effects and externalities of TFA, as well as some negative effects and externalities. Examining this evidence as a whole reinforces the need to develop a full scope of who benefits from TFA and who does not. It also motivates the need for an improved understanding of the mechanisms that support and undermine TFA's effectiveness across selection, training, and teaching. In doing so, it provides the foundation for my larger intellectual agenda, which involves examining how these various aspects of TFA interact to shape students' opportunities, and linking classroom observations of teaching with administrative data to begin disentangling how TFA teachers might differ from non-TFA teachers in their schools, and how this is related to student learning.

This thesis contributes to this larger agenda by examining treatment heterogeneity in TFA, attending to how the effects of TFA vary across the distribution, as well as whether having a TFA teacher in elementary, middle, and high school is differentially related to short and long term student outcomes. In doing so it provides policy relevant information about where TFA is successfully meeting students' needs, which grade levels are likely to have the largest payoffs, and where TFA has room for improvement.

[2] The annual growth rate equivalents noted here were calculated by the authors of the original Mathematica report. They are based on the Iowa Test of Basic Skills which is normed so that direct comparisons can be made across grade levels. While it would be useful to identify differences across grades, this is not possible given grade level variations in the standard deviations of the test being (see Bloom, Hill, Black, & Lipsey, (2008), for further discussion). For comparison, the mean effect size of 0.15 identified in math by Glazerman and colleagues falls somewhere between 15 and 25 percent of the average annual gains for grades 1-5 identified by Bloom and colleagues across a number of nationally normed tests that do not include the ITBS.

[3] It seems likely that this has been done within TFA, but non-affiliated researchers have not been able to do this thus far, but no external research does this currently.

[4] In 2013, TFA had an operating budget of nearly $200 million dollars, 30 percent of which came from public sources including the federal government and state and school district fees (Teach For America, 2013c).

[5] Teach For America prepares its teachers in one of several Summer Institutes, hosted in large partner districts across the country. TFA corps members gain classroom experience during Institute by teaching summer school under the supervision of a credentialed teacher from the partner district.

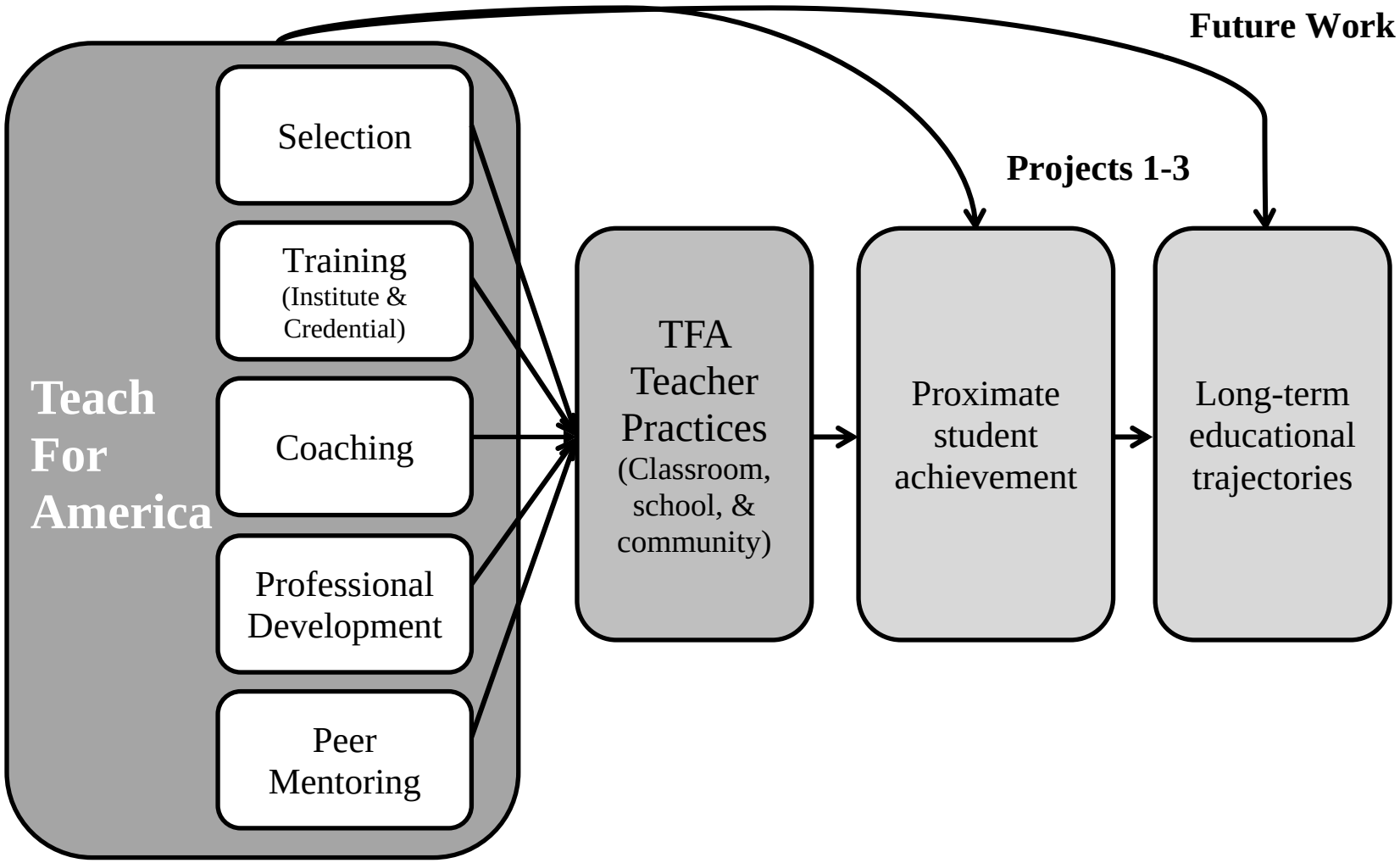


Figure 1.1. Conceptual Model of Teach For America

Chapter 2

Teaching for All?
Teach For America's Effects on the Distribution of Student Achievement

A mounting challenge facing America's educational system is the increasing gap between the educational achievement of students from poor and well-off families (Reardon, 2011) resulting from increases in school segregation, gaps in school quality, and educational spending across families of different income levels (Kornrich & Furstenberg, 2013; Reardon & Bischoff, 2011). As a consequence, low-income students have fewer high-quality teachers, teachers are more likely to leave low-income schools, and few teachers are interested in moving to low-income schools even when offered substantial pay increases to do so (Boyd, Lankford, Loeb, & Wyckoff, 2005a; Darling-Hammond & Sykes, 2003; Glazerman, Protik, Teh, Bruch, & Max, 2013; Hanushek, Kain, & Rivkin, 2004b; Isenberg et al., 2014). While school administrators have investigated a number of methods to improve teacher quality and supply in low-income schools (Ingersoll, 2004; Levin, 1968), districts increasingly rely on alternative pathways to teacher certification to recruit teachers. Of the growing numbers of alternative teacher preparation programs, Teach For America (TFA) is among the most influential.

TFA was founded on the simple, but unconventional premise that "good" teachers can be identified based on highly-selective background criteria, given limited amounts of training and ongoing coaching, and can be more successful in challenging, underserved schools than traditionally prepared teachers. TFA argues that its teachers can radically improve educational opportunities and achievement for low-income youth in underserved urban and rural communities across the US, with only a two-year commitment (Kopp, 2011). TFA's presence has expanded dramatically: In 2012, more than 10,000 TFA corps members taught 750,000

students in 46 regions across 36 states and the District of Columbia (Barahona, 2012). Lauded and critiqued for its impact on a variety of educational domains (Miner, 2010; Rotherham, 2011; Veltri, 2010), TFA judges its own success and failure based on student academic outcomes in its teachers' classrooms.

A growing body of work examines the average effect of TFA on student achievement, yielding contradictory results, with some finding that TFA teachers outperform their non-TFA counterparts (Glazerman et al., 2006) and others finding the opposite (Darling-Hammond et al., 2005). These mixed results suggest that TFA may not have a uniform impact for all types of students in all contexts, but might instead generate heterogeneous effects. Developmental science suggests that many types of interventions affect students differently (Duncan & Vandell, 2011), and existing examinations of TFA have only begun to consider how its effects might vary across different types of students and schools (Glazerman et al., 2006; Xu et al., 2011). I contribute to the growing literature on the impacts of TFA, and seek to inform larger conversations about teacher recruitment, training, and effectiveness in underserved communities by examining how the effect of TFA teachers varies across the distribution of student achievement. By focusing on goals related to closing achievement gaps and having students master grade-level content, TFA teachers might differentially boost achievement among some students than others. It thus becomes important to understand how TFA affects the distribution of achievement more broadly, and whether TFA teachers affect different parts of the distribution equally.

Prior Research on Teach For America

The relationship between TFA and student achievement has been examined using a variety of experimental, quasi-experimental, and descriptive research designs, in many grades and contexts. Two studies use randomized experiments to isolate the causal effect of TFA on

student achievement. Glazerman et al. (2006) evaluate the effect of TFA on student achievement in grades 1-5 using a random assignment design at six TFA sites, finding that students assigned to TFA teachers outperform students in novice and veteran non-TFA classrooms in mathematics, but not reading (ES = 0.15 SD). Clark et al. (2013) find that students randomly assigned to TFA secondary math teachers outperform students in comparison classrooms in 11 districts in eight states. TFA teachers outperform both alternatively and traditionally certified novice and veteran comparison teachers, with larger effects in high school than middle school.

In most studies, including those with the strongest quasi-experimental designs, TFA teachers have a positive effect on math and science, and little to no impact on language arts (Boyd et al., 2006; Henry et al., 2010; Kane et al., 2008; Xu et al., 2011). Most descriptive studies also find this pattern (Raymond et al., 2001; Strategic Data Project, 2012; Turner, Goodman, Adachi, Brite, & Decker, 2012; Ware et al., 2011). Across all of these studies, TFA teachers outperform novice teachers, and in some they also outperform veteran teachers. TFA teachers also outperform new teachers from selective undergraduate teacher preparation programs and teachers from teaching fellows programs with selective recruitment criteria (Boyd et al., 2012; Clark et al., 2013; Henry et al., 2010). Two descriptive studies find disconfirming evidence suggesting that TFA teachers outperform novice teachers, but are either statistically indistinguishable from (Noell & Gansle, 2009; Schoeneberger et al., 2009) or worse than certified veteran teachers (Darling-Hammond et al., 2005). Finally, one study finds that TFA teachers are less effective than veteran teachers and no different than novice teachers (Laczko-Kerr & Berliner, 2002).[6] On balance, the most rigorous evidence suggests TFA teachers outperform most novice teachers and some veterans, particularly in mathematics and science, but do not have a uniform effect across all subjects and students.

A Distributional Perspective on Teach For America

Distributional studies show that many education policy interventions do not have a uniform effect on all intervention recipients (Arulampalam, Naylor, & Smith, 2012; Lamarche, 2007). For example, findings from Project STAR suggest that class size matters most at the top of the achievement distribution (Jackson & Page, 2013). Given TFA's intent to promote achievement gains for all students and its expanding role in low-income districts and teacher policy, it is important to consider not just its average impact, but also the way it impacts the entire distribution of student achievement. We might expect several different distributional patterns for the effect of TFA, which may result from a number of TFA characteristics including the selective background of TFA teachers, their relative academic strengths, their commitment to equity issues, and their limited experience with students needing enrichment or intervention. Consider Glazerman et al.'s (2006) finding of a positive average effect of TFA teachers on math achievement. TFA teachers may be equally effective at improving student achievement at all points of the achievement distribution because TFA teachers' academic credentials, including test scores, coursework, and competitiveness of degree granting institution, surpass those of traditionally certified teachers (Boyd et al., 2012, 2006; Decker et al., 2004a), and provide them with greater depth of content knowledge which outweighs their relative lack of pedagogical experience. If academic background alone sets TFA teachers apart, then they may be equally more effective than non-TFA teachers across high, low, and average-achieving students. Given that TFA qualifications are especially strong in math and science (Xu et al., 2011), we would be most likely to see this pattern in those subjects.

Alternatively, TFA teachers' selective backgrounds might support learning for similarly higher-performing students, but they may lack the pedagogical content knowledge and curricular

knowledge needed to effectively teach struggling students (Shulman, 1986, 1987; Wilson et al., 1987), which Veltri (2010) suggests is a limited part of TFA training. If this were the case, the effect of having a TFA teacher might be restricted to the top of the distribution, promoting learning for some while simultaneously undermining equity.

But the opposite pattern could occur as well. If the cultural ethos of the TFA program leads corps members to develop a particular focus on equity, and if this mindset is reinforced with pressure exerted on corps members by TFA staff to raise all students' tests scores, in order to do so, TFA teachers would need to place particular efforts on low-performing students. TFA training and mentoring focuses heavily on data tracking of student performance, which may increase instructional support for low-performers among TFA teachers relative to non-TFA teachers (Goldstein, 2013). Goldstein (2013) suggests that TFA teachers are incredibly "mission driven and optimistic", and Dobbie and Fryer (2011) find that corps members have higher beliefs about the life chances of poor children than individuals who were nearly selected into the program, but not admitted. If the cultural ethos of TFA corps members and the intensive focus on improving test scores drives them to focus efforts on struggling students, then the effects of TFA teachers might be largest at the bottom of the distribution.

These scenarios are all plausible, and it is possible that these and other alternatives are occurring simultaneously. Given the variation in average effectiveness between reading and math in previous work, it is also possible that different patterns could hold across these two subjects, which may highlight that experience matters differentially for these subjects in elementary school. It is thus helpful to know how the entire distribution of student achievement is affected by having a TFA teacher. Examining variation in the impact of TFA can help to identify the relative strengths and weaknesses of the TFA program. Such findings will to speak to larger

debates about whether alternative certification programs, which recruit teachers from selective backgrounds, can be successful at reaching all students with limited training. They may also highlight areas of the TFA selection and training model which are not meeting students' needs. Finally, these findings might identify relative strengths of the TFA program, which with further study of the TFA training program, could yield insights for other teacher preparation programs serving low-income students.

Data

To understand whether the effect of TFA varies across the distribution of student achievement, this study estimates the quantile treatment effects (QTE) of being randomly assigned to a TFA versus non-TFA teacher. Mathematica Policy Research collected these data during the 2001-2002 and 2002-2003 school years, from six of the 15 active TFA regions, including Baltimore, Chicago, Los Angeles, Houston, New Orleans, and the Mississippi Delta (Decker, Mayer, & Glazerman, 2004b).[7] These six regions were randomly selected to represent the mix of districts served by TFA (predominately black vs. Hispanic, urban vs. rural), and within each region, schools with at least one TFA and one non-TFA teacher at the same grade level were chosen at random. The final sample included 100 classrooms, grades 1 to 5, at 17 schools, and a total of 1,969 students.

Student achievement is measured in the fall and spring using the Iowa Test of Basic Skills (ITBS). For my dependent variable, I use the Normal Curve Equivalent (NCE) math and reading scores, which are age-adjusted, and nationally normed to have a mean of 50 and a standard deviation of 21.06.[8] Because first graders took only a portion of the reading test, I exclude them from my analysis.[9] I further restrict my sample to students who have at least one fall and one spring test score. In addition, I treat previously unidentified invalid test scores (a raw

score of 99), given to 9.5 percent of 2nd through 5th graders in the public use data, as missing.[10] Descriptive statistics for TFA and non-TFA classrooms in the analytic sample are presented in Table 1.

[Insert Table 1 Here]

Method

To examine the effect of TFA on the distribution of student achievement, this paper estimates quantile treatment effects (QTE) (Firpo, 2007). QTE allow for unconditional comparisons of the achievement distributions of TFA and non-TFA students, and provide more information on the nature of treatment effects on the treated sample than mean differences. [11] In the context of experimental data, QTE are estimated by calculating the difference in the two marginal distributions (cumulative distribution functions, or CDFs) and are identified at each quantile in an analogous logic to average treatment effects under the potential outcomes framework.[12] Using these CDFs, I examine the difference between these two distributions at various percentiles of the outcome variable, ITBS reading or math test scores. For example, I estimate the QTE at the 0.50 quantile by subtracting the median test score of non-TFA students from the median test score of TFA students.

As an example, Figure 2.1 and Figure 2.2 Panel A show the CDFs and QTE for unweighted baseline math NCE scores. Figure 2.1 shows the CDFs for baseline math scores in TFA and non-TFA classrooms. The CDFs present math NCE scores on the x-axis with the cumulative percent of the sample on the y-axis. The horizontal distance between these CDFs at each point in the distribution, which equals the difference in NCE scores, is the quantile treatment effect at that percentile. Included on Figure 2.1 (and subsequent figures) are two vertical lines indicating the ITBS national mean (at 50 NCE points) and the mean for the non-

TFA classrooms (for un-weighted fall math scores this is 31.5 NCE points), which underscore that the majority of the sample scores below the national average.

Figure 2.2 Panel A shows the corresponding QTE for the CDFs shown in Figure 2.1, where the x-axis represents the cumulative percentiles of the distribution, and the y-axis represents the difference in NCE scores between TFA and non-TFA classrooms at each percentile. The score difference (solid line) is plotted along with pointwise 95 percent confidence intervals (dashed lines), which are calculated by stratifying on block and treatment status and bootstrapping the estimates 999 times. Figure 2.2 Panel A shows that most of the QTE point estimates are at or near zero for baseline math scores, except at the upper and lower tails. Between the 6^{th} and 10^{th} percentiles, there is a negative and significant difference between treatment and control where the confidence intervals do not include zero, suggesting some imbalance across the distribution in random assignment.

[Insert Figures 2.1 and 2.2 here]

Figure 2.2 Panel B mirrors Panel A, assessing the degree to which randomization successfully balanced fall scores across the distribution of reading achievement. Panel B shows that the differences between TFA and control classrooms are negative beginning above the 64^{th} percentile, and significant above the 90^{th} percentile. This suggests that randomization was even less successful for reading than math, with non-TFA classrooms having more higher-performing students at the outset.

To address the lack of balance on fall scores, I use an inverse propensity-score weighting approach as a nonparametric first step (Firpo, 2007), which allows me to balance baseline test scores across the two groups and to account for differences in the likelihood of being assigned to a TFA or non-TFA teacher in different grade levels and schools.[13] This also allows me to adjust

for differences in the presence of non-response and invalid test scores by including indicator variables for whether students had missing or invalid test scores. I first use a logistic regression model to predict assignment to a TFA or non-TFA teacher as a function of randomization block, baseline test score deciles for math and reading, whether the student had valid or invalid missing values for fall or spring scores, and the baseline demographic characteristics from Table 1. I calculate the predicted probability of being in the treatment group, $\hat{p}$, and construct weights of $1/\hat{p}$ for those in the treatment group and $1/(1-\hat{p})$ for the control group. As shown by the p-values for mean comparisons in Table 1, these weights balance the treatment and control groups on these observable dimensions. Further, a test of joint significance for these characteristics is not significant, suggesting that, when using the inverse propensity score weights, there are no differences across random assignment groups. Propensity score weighting allows me to obtain unconditional estimates while still adjusting for any post-randomization imbalance in baseline test scores, and missing values. Inverse propensity score weight adjusted fall QTE are shown in the Appendix A Figure A.2 and are described in greater detail there. Overall, the inverse propensity score weight adjusted QTE show balanced samples across the fall distributions in both reading and math, and I thus use the inverse propensity score weights to estimate the spring QTE for reading and math.

Results

The QTE results for spring math and reading tests are presented in Figure 2.3, Panels A and B. As with Figures 2.2 and 2.3, Figure 2.3 plots the test score differences between students in TFA and non-TFA classrooms (y-axis), for each percentile of the distribution (x-axis). When the solid line, representing the point estimate at a given quantile, is above zero students in TFA classrooms are out-scoring non-TFA students, and when the solid line is below zero students in

TFA classrooms are scoring lower than those in non-TFA classrooms. These differences are statistically significant when the area between the two dashed lines (representing the 95 percent confidence intervals) does not include the line marking zero on the y-axis.

In Panel A of Figure 2.3 the point estimates are positive across nearly the entire distribution of math, and for most of the distribution, the confidence intervals show that these differences are statistically significant at the 5 percent level. Although the point estimates are not the same across the distribution, and are as large as 6 NCE points at the 80^{th} and 88^{th} percentiles, tests comparing differences between various percentiles across the distribution suggest that they are not statistically different from one another at the 5 percent level. Even though we cannot rule out an effect of zero for some portions of the distribution, we can largely rule out a negative effect of TFA on math achievement, except at the very upper tail. Thus, writ large, TFA's effect on math can be characterized as positive on average and shared throughout most of the distribution, though in a few parts of the distribution it might be more accurate to characterize it as non-negative.

[Insert Figure 2.3 Here]

Figure 2.3 Panel B shows the QTE for spring reading scores. Here the QTE plot suggests variation in the effect of TFA across the distribution. At the lower tail, the point estimates of the effect of TFA are negative, ranging from negative 2-4 NCE points. This effect is not statistically significant at the 5 percent level, but a further examination of the 999 replicates to examine what fraction of the replicates provide estimates which are positive, negative or zero for these quantiles provides additional evidence in support of this pattern. Within the quantiles 2-36 where the negative point estimates are found, we typically find that among the replicates used for producing confidence intervals, 0-2 percent of these group differences are positive,

approximately 5-10 percent are zero, and 90 to 95 percent are negative.[14] On balance, this suggests that we can rule out positive effects on the bottom of the distribution, and that the effects are likely negative.

In contrast, above the 40th percentile the point estimates are positive, ranging from 2-4 NCE points. Here again most point estimates are not significant with 95 percent confidence intervals, but there is suggestive evidence of this pattern when using 90 percent confidence intervals. With a few exceptions the confidence intervals indicate that the effect of TFA on this part of the distribution is not negative, and in many places is positive. Comparisons between percentiles at the bottom and the top of the distribution indicate significant differences between the effects of TFA on these portions of the distribution. Combined, these countervailing patterns of results yield an average effect that is near zero and statistically insignificant.

In addition to comparing TFA teachers to all non-TFA teachers, it is also important to determine whether TFA performance is similar to that of veteran non-TFA teachers. Results comparing TFA teachers to non-TFA teachers with more than three years of experience are shown in Figure 2.4.

[Insert Figure 2.4 Here]

The results shown in Figure 2.4, Panels A and B indicate that the pattern observed when comparing TFA teachers to all non-TFA teachers is more pronounced when TFA teachers are compared to non-TFA teachers with more than three years of experience. The QTE graphs in reading and math have the same shape as in the full-sample comparisons. For math, the pattern is nearly identical, but the confidence intervals are tighter. Likewise, in reading the negative treatment effect at the bottom of the distribution is statistically significant at the 5 percent level for half of the percentiles examined, as is the positive effect at the top of the distribution.

Additional comparisons (not shown) relative to novice non-TFA teachers are less precise, given that the overwhelming majority of the comparison teachers are experienced, however the same general pattern of effects holds.

Discussion

This paper extends prior research on Teach For America by examining how student achievement in TFA and non-TFA classrooms varies across the distribution of achievement. It identifies variation in effects that was previously hidden by examining only average impacts of TFA. The distributional findings reveal different patterns for math and reading.

In math, students assigned to TFA teachers outperform control students throughout most of the distribution. This is consistent with overall positive effects of TFA on math observed in Glazerman et al. (2006), and evinces the fairly wide-spread effectiveness of TFA teachers relative to same-school comparison teachers. The magnitude of the differences is striking: the largest TFA effect of 6 points corresponds to .28 SD of the nationally normed sample and .34 SD of the control group's fall score.[15] This effect corresponds to approximately three months of instruction, falling somewhere between the .22 SD effect observed in the Tennessee STAR class size experiment (Krueger, 1999), and the 0.35 SD effect of KIPP schools (Angrist et al., 2010).

By contrast, the pattern of effects in reading varies across the distribution. TFA teachers appear to have a positive impact on the top of the reading distribution that is as large as .19 SD, even relative to veteran teachers. However, TFA students at the bottom of the reading achievement distribution are scoring worse than their counterparts in non-TFA classrooms, particularly compared to peers in experienced teacher classrooms. The null mean effect for reading found in previous research thus appears to be concealing important distributional differences.

When contextualizing these findings is important to consider that the distributions presented here are only comparable to schools similar to those in which TFA teachers are placed. As Figure 2.1 highlights, although some students among this sample perform above the national average, the national mean represents approximately the 87th percentile for this sample, while the mean score is roughly 31 NCE points for both reading and math, nearly one standard deviation below the national average. Thus, the results cannot be generalized to schools that are in higher-performing neighborhoods and districts with no TFA presence.

It is also important to consider what learning about distributional variation in the impact of TFA can and cannot tell us about the costs and benefits of alternatively certified teachers, teacher selection, training, and experience. The TFA model makes it impossible to disentangle whether the changes in student learning resulting from TFA are a product of TFA's selection, its training, or some combination thereof. Existing research demonstrates that selection matters, as students learn more under TFA teachers who score highly on some TFA selection criteria in the application process (Dobbie, 2011). However, training is also important, as teachers from other selective teaching fellows programs are, in some cases, less effective than traditionally certified teachers (Clark et al., 2013). To the degree that selectivity accounts for TFA's success, this suggests that policies should better incentivize teaching. Further, TFA training is not an entirely uniform intervention. To the degree that training is also important, more work is needed to identify the specific facets of teacher training and ongoing support that contribute to the success of TFA teachers. While the TFA intervention is not entirely uniform (due to regional variation in summer training and local teacher preparation programs), this variation could be helpful in identifying particularly important aspects of TFA training. Given the compressed timeframe in which TFA training occurs, one might hope that any such insights could be incorporated into

existing training for new teachers or professional development for current teachers without requiring an extraordinary commitment on behalf of teachers.

In addition, there are important aspects of instructional support for low-performing readers that TFA must strive to improve. Selective backgrounds and minimal training are insufficient for supporting low-performing readers and TFA should consider ways to address literacy instruction to support such students. Previous research uses the cutoff of the 34th percentile to help identify students at-risk of special education services (Woodward & Baxter, 1997). This suggests that many of the students who are underperforming in TFA classrooms might particularly benefit from targeted reading interventions requiring specialized training or experience. Thus, particular attention to the specific instructional practices and use of institutional resources that non-TFA teachers employ to support their lowest-performing readers seems warranted. Further, future work examining how TFA and non-TFA classrooms differ in pedagogy, classroom organization, and targeted intervention might provide valuable insights into how teachers from different backgrounds can work together to meet the needs of our underserved students.

In sum, when evaluating TFA teachers by TFA's own rubric—student achievement— my results suggest that both proponents and detractors are partially right: TFA clearly raises math scores throughout the distribution, while in reading it appears to raise scores for high achievers and lower scores for low achievers, particularly relative to veteran teachers. Thus, TFA does improve student achievement for some students in low-income schools in some subjects, but must further refine its model to meet this goal for all students.

[6] In some of these studies, TFA teachers are compared to teachers in schools or districts that do not have TFA teachers, extending beyond the counterfactual teachers in TFA schools and districts. One such example, Noell

and Gansle (2009), compares TFA teachers to teachers across the state, finding positive impacts compared to novices and no differences with veterans. While this type of comparison is potentially informative, it also may confound issues of selection across school types such that any positive or negative effects of TFA teachers cannot be separated from the sorting of students into different types of schools. In contrast, a stronger research design includes school fixed effects and only compares TFA teachers with non-TFA teachers in the same schools because these are the only types of teachers that most low-income students have access to and are therefore the best counterfactual, for example Boyd et al. (2010) or Henry et al. (2010).

[7] The pilot region, Baltimore, was studied during the 2001-2002 school year. The other five regions were studied from 2002-2003.

[8] Normal Curve Equivalent scores are calculated from raw scores which are then normed based on grade and quarter of the school year (fall, winter, or spring) using the national ITBS sample, and converted into rankings such that the distribution of scores is normal. This allows for cross age and cross-grade comparisons of scores at equal intervals.

[9] While students in grades 2-5 received reading scores that were calculated using responses from tests of both vocabulary and word analysis, first grade reading tests were scored separately as vocabulary and word analysis, and no combined score is available (Glazerman & Grinder, 2004). To facilitate comparison across reading and mathematics achievement, I likewise only examine mathematics achievement among students in grades 2-5. Results presented are robust to the inclusion of first graders; in these analyses I use students' vocabulary scores to match Glazerman et al. (2006).

[10] Below the score of 99, the next highest raw score observed was 41, while the highest possible raw score at any grade level is 44 in reading and 50 in math (Hoover, Dunbar, & Frisbie, 2007; Riverside Publishing, 2012). Raw scores of 99 corresponded to NCE scores of 0. Communication with Riverside Publishing confirmed that 99 is an invalid raw score, and that 0 is an invalid NCE score.

[11] Concurrent with this paper, which was presented in February 2013 at the annual Sociology of Education Association meeting and in March at the spring meeting for the Society for Research on Educational Effectiveness, a working paper released by Antecol, Eren, & Ozbeklik (2013) in April also examines distributional differences in the effects of TFA on student achievement in elementary school using the MPR data. While I use inverse propensity score weights to calculate QTE, Antecol et al. use fixed effects quantile regression models which report differences at conditional quantiles (see Firpo, Fortin, and Lemieux (2009) for a discussion of the shortcomings of conditional quantiles). See the online supplement for further discussion of the distinctions between this paper and Antecol et al. (2013).

[12] For more details, see the online appendix.

[13] Glazerman et al. (2006) include a sample normalization weight in their estimates to account for the fact that each block has a different number of TFA and non-TFA classrooms, with slightly different numbers of students in each classroom, making the odds of assignment to treatment non-uniform. This variation is accounted for by including block fixed effects in the inverse propensity score weights.

[14] See the online appendix for more details.

[15] The national sample has a SD of 21.06, whereas the universe of schools served by TFA includes few students scoring at the top of the national distribution, resulting in both a lower mean and a smaller SD in this sample (SD=17.69). Thus, for understanding how much students benefit from having a TFA teacher relative to other students in their schools, the latter is the appropriate effect size, but the former provides information on the effect relative to the national population of students. Glazerman and colleagues use the sample SD to calculate their average treatment effect estimates of 0.15 SD in math and 0.03 SD in reading.

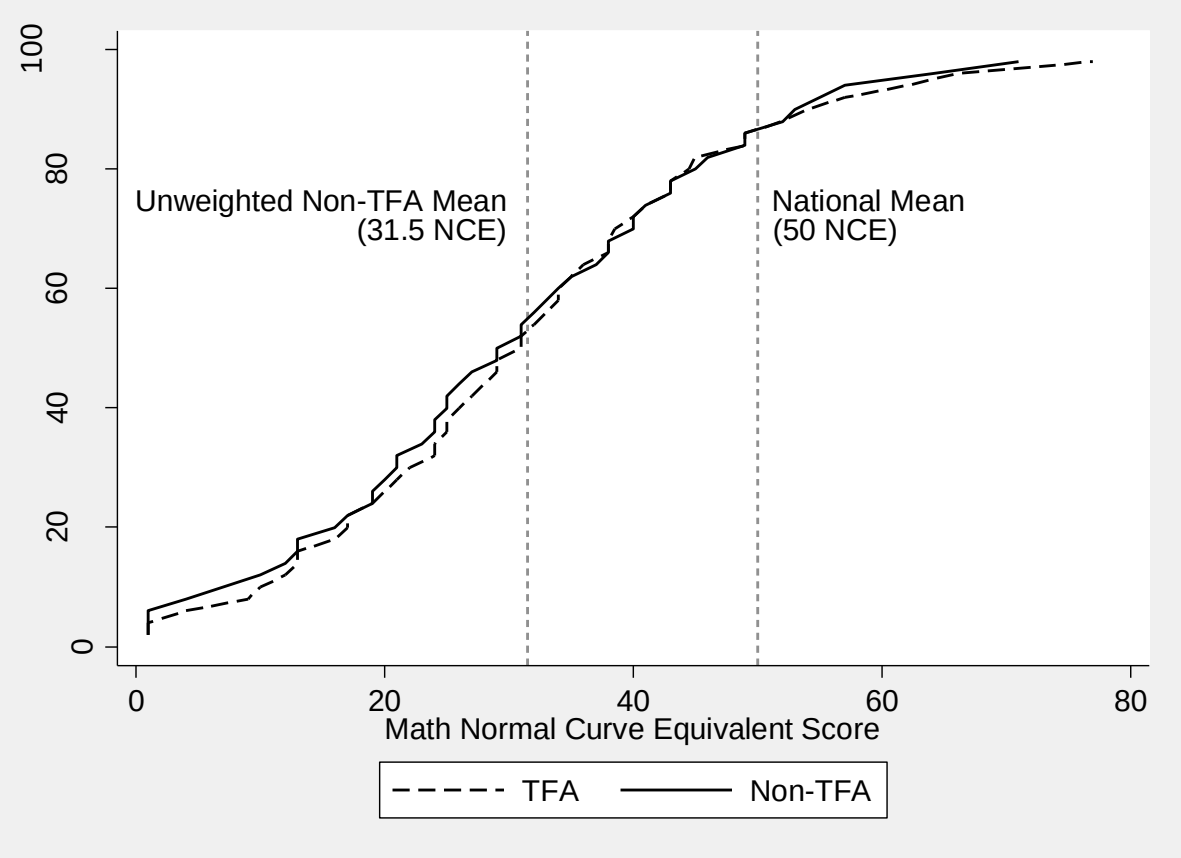
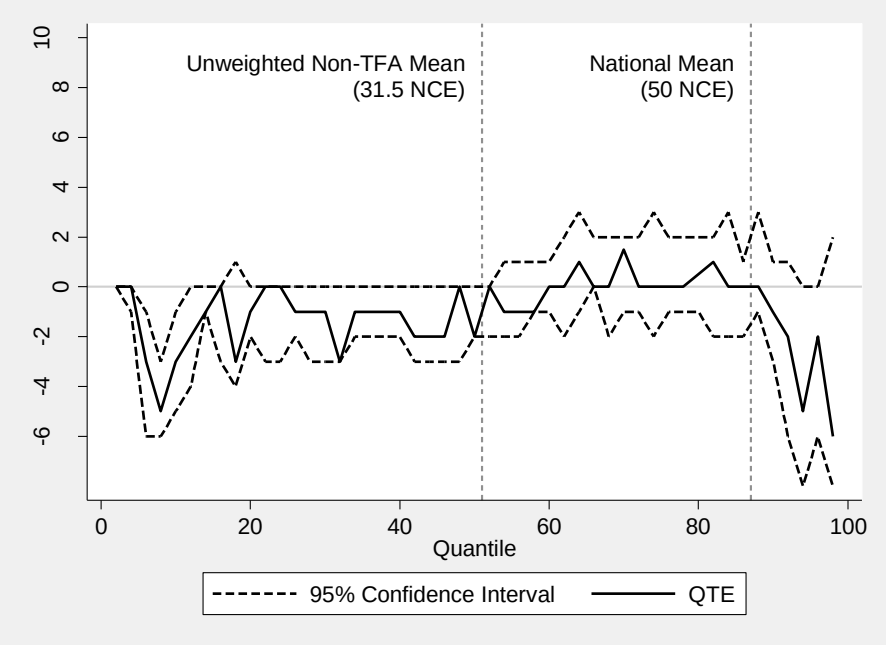


Figure 2.1. Cumulative Density Functions (CDFs) for fall math achievement in TFA and non-TFA classrooms.

Notes: Figure shows cumulative distribution functions for baseline math Normal Curve Equivalent scores from the Iowa Test of Basic Skills separately for TFA classrooms and non-TFA classrooms. Estimates are unweighted. Data from the Mathematica Policy Research Teach For America Evaluation.

Panel A. Unweighted differences in fall mathematics achievement



Panel B. Unweighted differences in fall reading achievement

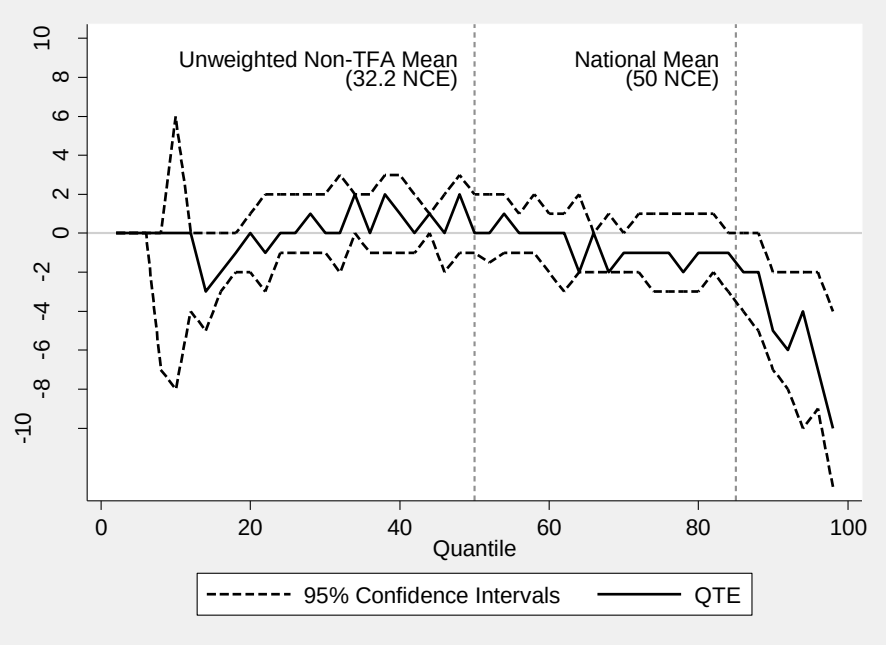
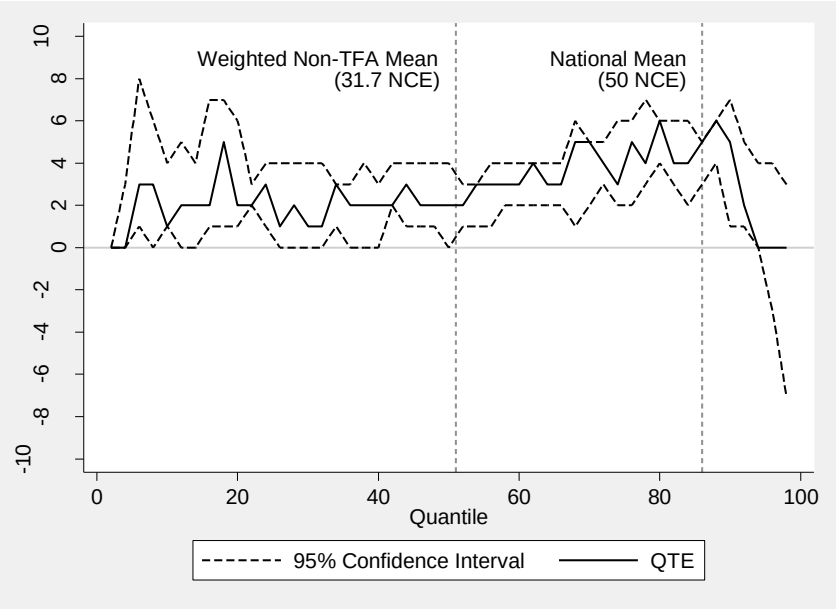


Figure 2.2. Unweighted quantile treatment effect estimates of the impact of assignment to TFA classroom on reading Normal Curve Equivalent scores at baseline (fall).

Notes: Panels A & B of figure show QTE for the effect of being assigned to a TFA classroom on math and reading Normal Curve Equivalent scores from the Iowa Test of Basic Skills at baseline. Estimates are unweighted. Data from the Mathematica Policy Research Teach For America Evaluation.

Panel A. Weighted differences in spring mathematics achievement



Panel B. Weighted differences in spring reading achievement

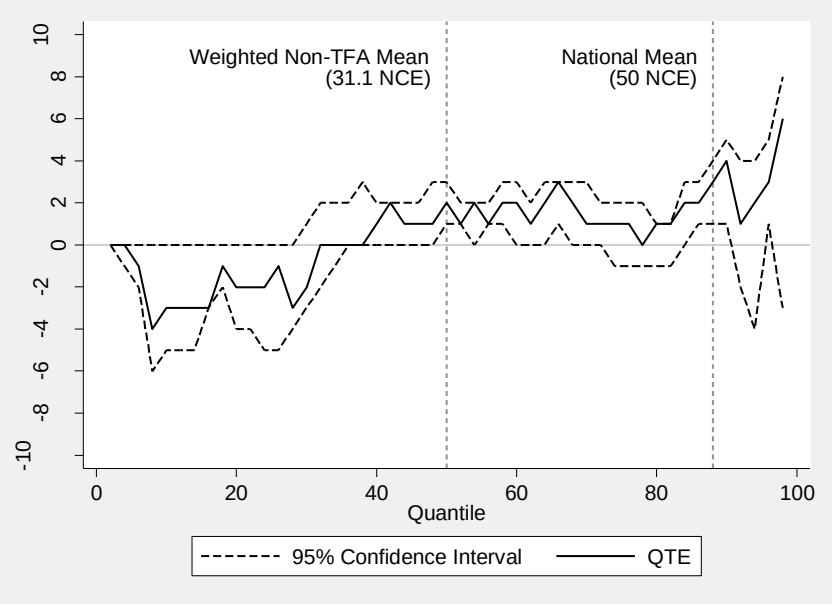
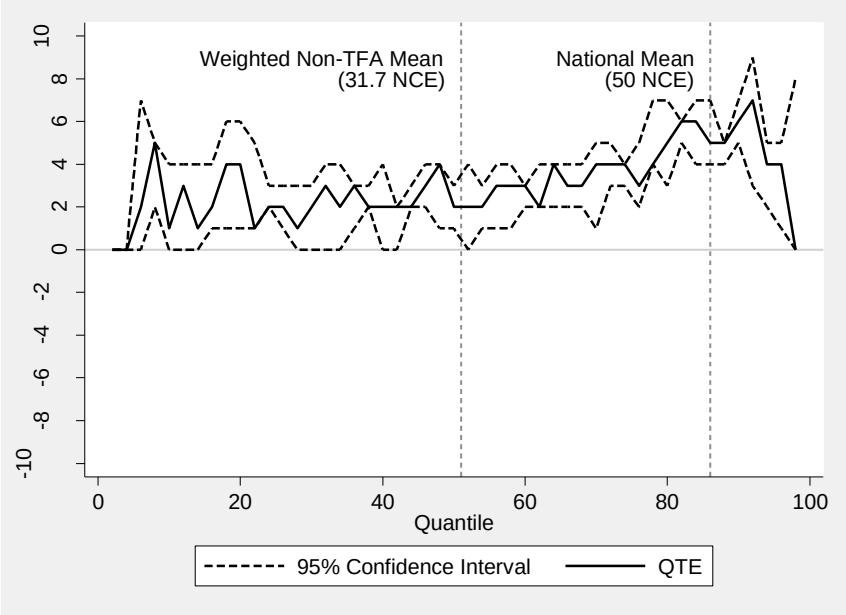


Figure 2.3. Inverse propensity score weighted quantile treatment effect of assignment to TFA classrooms on posttest (spring) test scores, TFA vs. all non-TFA teachers.

Notes: Panels A & B of figure show QTE for the effect of being assigned to a TFA classroom on math and reading Normal Curve Equivalent scores from the Iowa Test of Basic Skills in the spring following random assignment. Estimates are weighted using inverse propensity score weights. Weights are $1/\hat{p}$ for treatment observations and $1/(1-\hat{p})$ for control observations, where $\hat{p}$ is generated from a logistic regression of treatment status on baseline demographics, sample design variables, and baseline test score deciles. 95% CIs are obtained by bootstrapping with replacement within randomization block. Data from the Mathematica Policy Research Teach For America Evaluation.

Panel A. Weighted differences in spring mathematics achievement



Panel B. Weighted differences in spring reading achievement

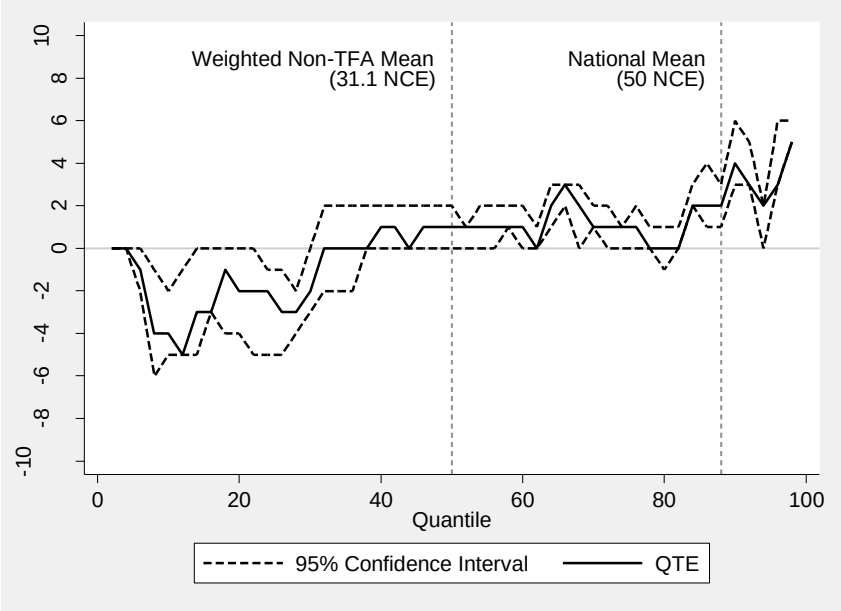


Figure 2.4. Inverse propensity score weighted quantile treatment effect of assignment to TFA classrooms on spring test scores, TFA vs. veteran non-TFA teachers only.

Notes: Panels A & B of figure show QTE for the effect of being assigned to a TFA classroom on math and reading Normal Curve Equivalent scores from the Iowa Test of Basic Skills in the spring following random assignment relative to veteran teachers. Estimates are weighted using inverse propensity score weights. Weights are $1/\hat{p}$ for treatment observations and $1/(1-\hat{p})$ for control observations, where $\hat{p}$ is generated from a logistic regression of treatment status on baseline demographics, sample design variables, and baseline test score deciles. 95% CIs are obtained by bootstrapping with replacement within randomization block. Data from the Mathematica Policy Research Teach For America Evaluation. Veteran teachers have 4 or more years of experience.

Table 2.1.
Baseline characteristics of study sample and missing values[1]

	Control Mean	T-C Difference	SE	P-value[2]
Female	0.474	0.018	0.018	0.319
Black	0.720	-0.008	0.034	0.809
Hispanic	0.238	0.016	0.021	0.459
Over age for grade	0.231	0.018	0.024	0.447
Free/reduced lunch eligible	0.975	-0.002	0.009	0.786
Did not move classes during school year (stayer)	0.928	0.018	0.014	0.195
Percent of class in research sample at end of year	0.817	-0.018	0.030	0.556
Math pretest Normal Curve Equivalent score	30.638	-0.177	0.840	0.834
Reading pretest Normal Curve Equivalent score	31.141	0.245	0.832	0.771
Sample size:	N = 1430			
Joint test for baseline child characteristics:	p = 0.224			

[1] Sample includes 2nd - 5th graders with at least one valid pre- and post-test

[2] P-values calculated after using inverse propensity score weights and clustering on randomization block

Chapter 3

Stage-Proficiency-Environment Fit and the Effectiveness of Teach For America

Introduction

Child development research underscores the importance of developmentally appropriate interventions, and suggests that the intervention effects might vary across different stages of development (c.f. Duncan & Vandell, 2011). Teach For America (TFA) is an alternative teacher certification program that recruits recent graduates and young professionals to teach in low-income schools for two years. It is one of the highest-profile interventions targeting teaching and learning in low-income schools, and yet most existing evaluations of the program fail to consider the possibility of variation in its effects stemming from the developmental fit between the program and the students it serves. Only two of the current evaluations of the TFA program even examine its effects across the range of grades to which TFA teachers are assigned.

This chapter considers how the relationship between having a TFA teacher and student achievement might vary across elementary, middle, and high school. Building from Eccles' and Midgley's (1989) theories about stage-environment fit, I examine whether TFA's effects are largest in a particular developmental stage, as identified here by school grade. I also examine the notion of the match between the TFA program and the students it serves, by considering whether TFA is more effective for students with higher or lower levels of baseline proficiency. I test whether the effects of TFA are more in line with an *accumulated advantages* or *skill-begets-skill* hypothesis (Cunha et al., 2006; 2007), or whether TFA functions in a more *compensatory* manner (c.f. Morgan, Farkas, & Hibel, 2008; Rutter, 1987; Sameroff & Chandler, 1975). Modifying the theoretical frame of stage-environment fit (Eccles & Midgley, 1989), I test for

variation in the effects of TFA based on the stage-proficiency-environment fit. I also examine whether these patterns are consistent across subject area domains.

My results suggest that, relative to conventionally trained teachers, TFA teachers improve student learning across all three developmental stages, with the largest effects found in high school. I also find evidence of variation in the effects of TFA across academic domains, showing that TFA teachers are most effective in math and science, particularly in high school. Interestingly, I also find some evidence of negative effects on elementary reading, a finding that echoes the results presented in Chapter 2. Finally, I show that the effects of TFA also vary by students' baseline proficiency levels with higher-skill students benefitting the most in elementary and middle school and lower-skill students benefitting the most in high school. Together, these findings suggest that TFA's effects are far from uniform, but are nonetheless positive in most areas, save elementary reading. These results also underscore the need to test for impact variation and to investigate how a program is affected by its match with the developmental stage and entering skill level of the children it hopes to serve.

Literature Review

Many of the most noteworthy educational interventions and innovations in recent decades target resources at a particular age group. Successful programs like Head Start and Abecedarian focus on early childhood, the Tennessee Star experiment randomized students into classes of different size in early elementary, and the AVID program supports teens in their transition to college. While evidence suggests that larger benefits can be gained by concentrating resources on a narrow age range, particularly early childhood, the opportunity gap for many low-income children is such that participation in early childhood programs is not universal and interventions are needed throughout many children's academic careers. Few large-scale educational

interventions target children across the range of developmental stages encompassed by K-12 education, and many of those that do, such as the Harlem Children's Zone or school choice vouchers, are only available for students based in a particular location.

In contrast, Teach For America (TFA) is somewhat unique in that it provides a year-long intervention to children ranging from pre-K through 12th grade in 52 regions throughout the country. Although TFA teacher training has many commonalities across different school levels, it is unclear whether this training produces effects of similar magnitude for students at different developmental stages. Many factors, such as differing amounts of content knowledge and differences in school context and structure, could contribute to differential effectiveness across grade levels. However, most evaluations of the impact of TFA on student learning examine it as though it were a fairly uniform intervention.

In contrast, I argue that TFA is a complex intervention adapted to many different types of school contexts to match the developmental stage of the students it serves. Although TFA trains teachers to work in elementary, middle, and high schools, and provides largely separate, but culturally similar training for each school type, prior research rarely distinguishes between the effects of having a TFA teacher at each of these different developmental stages. As a result, the relationship between having a TFA teacher and student achievement is seldom investigated across school levels.

Further, the only two studies to date that examine TFA teachers across elementary, middle, and high school do not test whether the impact of TFA teachers varies across these levels (Henry et al., 2010; Ware et al., 2011). Thus, while we might expect variation in the effects of having a TFA teacher on student achievement across grade levels, no research has investigated whether these effects are statistically different. As variation in the effectiveness of TFA teachers

by school level is an inherently interesting question to district leadership partnering with TFA to employ corps members, and to policy makers more broadly, this paper examines whether TFA has heterogeneous impacts across these stages.

Examining the effects of TFA across developmental stages is particularly important for underserved school districts seeking to make the most of their limited resources. Recent estimates suggest that districts pay between two and five thousand dollars per corps member they recruit, making them more expensive than traditionally certified teachers(Simon, 2012).[16] To best allocate scarce teacher-recruitment resources, it is important to understand where TFA is most effective. It is also important to identify the school levels and subjects for which TFA is not effective and investigate potential mechanisms for why TFA is ineffective and what might be done to improve the effectiveness of TFA teachers in the identified age(s) and subject matter(s). This chapter fills this gap in the literature by examining whether TFA has a differential relationship with student achievement across elementary, middle, and high school.

The literature examining the potential mechanisms by which TFA might differentially impact student achievement across developmental stages is reviewed in Chapter 1. Here I provide a brief summary.

Research on TFA across Developmental Stages

Previous research examining the relationship between TFA and student achievement, typically focuses on either elementary, middle, or high school, and few studies examine TFA's effects across all three school levels. Research examining the effects of TFA in elementary school is mixed, with some finding positive effects in math and no effects or negative effects in language arts (Darling-Hammond et al., 2005; Glazerman et al., 2006). Studies focusing on middle school or elementary and middle school together find that students perform better in TFA

classrooms in math, but results in language arts range from positive to neutral to negative (Boyd et al., 2010, 2006; Noell & Gansle, 2009; Strategic Data Project, 2012). Research in high school also indicates positive effects in science, and math, with mixed results in language arts (Schoeneberger et al., 2009; Xu et al., 2011). Only Henry et al. (2010), and Ware et al. (2011) have extended their analyses to examine impacts in all school types, finding evidence of positive effects, particularly in math and science in high school, with some smaller effects in language arts in some grades, but neither make explicit comparisons across these three levels.

The present study builds directly on Henry et al. (2010), who examined the impact of teachers on student achievement in elementary, middle, and high school grades in North Carolina in 2004/05 through 2007/08 trained through a variety of preparation pathways, comparing them relative to University of North Carolina (UNC) trained teachers. Most notably, Henry and colleagues find that TFA teachers outperform teachers trained through UNC pathways in high school math, English, and science, as well as middle school math. However, they find no differences in elementary math or reading, middle school reading, or high school social studies relative to UNC trained teachers.

Henry et al.'s work suggests that differences in the relationship between TFA and student achievement across developmental stages are present in North Carolina. My study builds on this work in several ways. First, rather than comparing TFA teachers relative to University of North Carolina-trained teachers, it compares TFA teachers with all other types of teachers in the same schools, grades, and years. The TFA effects relative to all North Carolina teachers might differ from effects relative to exclusively UNC trained teachers, and would likely be larger given that the UNC training program is relatively selective compared to many other traditional teacher preparation programs. Second, I test for differences in the effects of TFA across developmental

stages, which has not been investigated by prior research on TFA. Finally, I examine not only the developmental fit of the TFA program, but also whether TFA has differential effects based on the match between students' baseline proficiency, which was also not a part of the Henry et al. investigation.

Theorizing Heterogeneous Treatment Effects

The impact of TFA on student achievement might vary across school types for several reasons related to variation in the quality of the TFA teachers and their training, as well as the quality of the counterfactual non-TFA teachers in the same schools. Many of these potential mechanisms were described in Chapter 1. Here I briefly review this research, arguing that the degree to which TFA "fits" as an intervention could vary across 1) developmental timing (i.e. across elementary, middle, and high school); 2) subject area (e.g. math versus English language arts); and 3) student baseline proficiency.

Developmental timing. Duncan and Vandell (2011) argue that it is important to consider the developmental timing of an intervention and argue that particular aspects of children's normative development can contribute to differences in program effects in meaningful ways. They point to earlier work by Morris, Duncan, & Clark-Kauffman (2005) who found treatment variation in achievement outcomes across children of different ages from interventions aimed at increasing parental income. While younger children benefitted from the intervention, older children were often tasked with additional childcare responsibilities and as a result, their achievement suffered. This example, and others like it in the literature (c.f. Duncan & Vandell, 2011), illustrate the need to examine whether interventions intended for children across multiple developmental stages have uniform effects.

Research examining a wide variety of interventions suggests that the effects of an intervention like TFA might be most effective across the developmental stages that roughly correspond to elementary, middle, and high school. While these school levels do not perfectly conform to the developmental stages of middle childhood, adolescence, and late adolescence/early adulthood, transitions across these school levels in the United States represent major transitions and developmental milestones, even if their timing does not correspond to different developmental periods for every child. They also represent transitions in terms of the social organization of schools, the social and academic expectations placed on children, and the level of social support provided by teachers versus peers, which also vary in the degree to which they complement the developmental needs of children. As a result, I operationalize developmental stages along the lines of the three typical school levels of elementary, middle, and high school.

Research examining the effectiveness of earlier versus later interventions provides some evidence that elementary school might be the ideal time to maximize the program's effects. Several school-based interventions have typically found larger effects in earlier grades. For example, Finn and Achilles (1999) find larger impacts of the Tennessee Star class-size experiment in Kindergarten and first grade than in later grades. Likewise, intensive reading interventions through the Response to Intervention (RtI), which have been shown to be effective in elementary grades, showed disappointing results when tested in a middle school experiment because it was difficult to implement the program as designed (Fuchs, Fuchs, & Compton, 2010). Similarly, a meta-analysis of experiments evaluating the effects of career education interventions finds larger effect of such programs targeted at elementary, rather than middle or high school students (Evans & Burck, 1992).

In addition to experimental evidence about the relative success of interventions aimed during the elementary school period, theoretical and observational research examining the structure and organization of schools also suggests that the developmental supports found in elementary school might make interventions focused during this time period more effective, in part because the exposure of students to a single teacher (i.e. dosage) is greatest in elementary school. Research provides findings congruent with a dosage framework, as interventions aimed at extending the school year and the school day suggest that increased school time has positive impacts on student achievement (Bellei, 2009; Skandera, 2012), and Lee et al. (2006) find that students learn more in full-day kindergarten than in half-day kindergarten. Research examining variation in time allotted to elementary school instruction in a given domain, such as math, provides similar findings, linking instructional time in a domain to achievement in that domain (Brown & Saks, 1986; Entwisle & Alexander, 1999). Thus, TFA teachers with more exposure to a given student should have a greater impact.

One way that this increased exposure might operate is if self-contained classrooms allow effective teachers to integrate the curriculum across the school day, reinforcing concepts across subject areas and providing ongoing emotional and academic support. Developmental science suggests that children learn best in settings in which curricula are integrated and subject-matter is interrelated (Bredekamp & Rosegrant, 1992; Cobb, 1994; Dewey, 2004; Piaget & Inhelder, 1969). Self-contained classrooms are a better developmental fit for children and adolescents, because a single teacher is responsible for instruction across subjects, can better meet the emotional, behavioral, and academic needs of students, which supports academic self-concept, motivation, school persistence, and student achievement (Eccles & Midgley, 1989; Eccles et al., 1993, 1996). As outlined by the stage-environment fit hypothesis, Eccles and Midgley (1989)

argue that the organizational structure of middle school, in which students change classes for different subjects multiple times per day, is mis-matched with adolescents' needs for increased autonomy and supportive teacher-student relationships. Traditional middle schools are thus a bad fit for adolescents because they are characterized by weaker teacher-student relationships that provide less autonomy for students and lead teachers to feel less effective instructionally (Brophy & Evertson, 1978; Eccles & Midgley, 1989; Midgley et al., 1989), and the transition to this organizational structure is particularly ill-timed with the developmental transition to adolescence (Eccles et al., 1993), and has negative impact on student achievement (Alspaugh & Harting, 1995).

School structures that are better-aligned with students' developmental needs, such as elementary and K-8 schools, should have larger effects, in part because TFA teachers can take advantage of the supportive environment to focus more energy on developing high quality instruction and less on addressing organizational and behavioral challenges. Prominent examples of this type of environment are the KIPP schools, which provide a highly structured environment with increased exposure to high-quality teachers. Teachers in KIPP schools can devote substantial amounts of time to supporting student learning, allowing effective teachers to have substantial impacts on student outcomes (Angrist et al., 2010). If TFA teachers are better able to take advantage of developmentally aligned school environments, then their impact is likely to be largest in the settings in which they have more contact with students and are able to be the most supported by the school environment. This is most likely to occur in elementary school, and less likely in traditional middle schools and high schools.

By contrast, it may be the case that because the middle school structure is such a poor fit, a highly motivated, so called "Superman" teacher can make a large difference (Gudmundsdottir

& Saabar, 1991; Spencer, 2012). This logic suggests that rather than TFA teachers being least effective where structures are generally not well-aligned with students' needs, this lack of alignment allows for teachers to play a particularly important role, in part because the other support for students is so weak. Eccles et al. (1993) find that middle school teachers have particularly low feelings of self-efficacy, which has been found to be related to poor student motivation and achievement. This suggests that counterfactual middle-school teachers may be especially ineffective, allowing for especially large impacts in middle school grades if TFA teachers can overcome middle school's disadvantages and provide their students with enriching experiences. In addition, evidence suggests that very high quality middle-school interventions can positively impact student outcomes. Dobbie and Fryer (2013) find that providing access to high-quality charter schools, like the Harlem Children's Zone Promise Academy, even as late as middle school, can positively impact college attendance and math achievement. If TFA teachers can have even a portion of the impact that the more comprehensive HCZ program has, this theory would suggest that middle school is an especially good time to leverage such an intervention because the alternative is so bleak.

Finally, others instead argue for special intervention efforts aimed at high school students, as delinquency and dropout problems persist and many students do not complete high school even after having satisfied the requirements to enter a four-year college (Darling-Hammond, 2006, 2007; Roscigno, 1999; Rumberger, 1987; Suh, Suh, & Houston, 2007). While people adhering to this perspective also value intervention at earlier ages, they argue that we need to make efforts to support struggling high school students to prepare them for the labor force and college (Hawkins, Catalano, Kosterman, Abbott, & Hill, 1999; Walker & Shinn, 2002). In particular, they point to the need for additional support to counter the high costs of student

dropouts in terms of labor market success, income, unemployment, crime, and incarceration (Murnane, 2013; Orfield, Losen, Wald, & Swanson, 2004), and highlight that even earning GEDs leave people at an economic and employment disadvantage relative to high school graduates (Cameron & Heckman, 1993; Rumberger & Lamb, 2003).

Limited evidence suggests that high school interventions and teachers can address issues like increasing high school completion and college going. Murnane (2013) provides an excellent summary of several successful interventions, including the Talent Development High School Model, which provided incoming 9th graders with learning communities and New York City's replacement of large high schools with small schools (H. S. Bloom, Thompson, & Unterman, 2010; H. S. Bloom & Unterman, 2012; Kemple, Herlihy, & Smith, 2005). Croninger and Lee (2001) find that high school teachers are an important source of social capital, and can help to reduce dropout rates by up to 50 percent. Erickson, McDonald, and Elder (2009) find that low-income students are the least likely to have adult mentors, but when they do, mentors, particularly teacher-mentors, have a large *compensatory* impact on attainment. Research from Stanton-Salazar and Spina (2003) suggests that even temporary assistance and support from mentors in high school can have long-term consequences for developmental trajectories. Finally, an experiment providing high school students with mentoring and cash incentives found increases in college going of 15 percent, and that cash incentives without mentoring had no impact (Carrell & Sacerdote, 2013). This research suggests that high school teachers can play an important role in mentoring students. Given the emphasis TFA places on fostering student-teacher relationships, we might expect high school TFA teachers to have particularly large effects on their students.

An additional factor which may also contribute to TFA's effectiveness in high school is teacher content knowledge, especially in subjects like mathematics and science. Prior research indicates that teachers' subject-matter content knowledge is an important predictor of student achievement. Several studies identify a positive relationship between teacher subject-matter knowledge or competence, as measured by teacher certification exams, and student achievement (e.g. Boardman et al., 1977; Ferguson, 1991; Greenwald et al., 1996; Hanushek, 1972; Harbison & Hanushek, 1992; Mullens et al., 1996; Rowan et al., 1997; Strauss & Sawyer, 1986; Wayne & Youngs, 2003). In North Carolina in particular, this positive relationship between subject matter measured by teacher certification exams is positive and significant for math and science, but negative for English (Clotfelter et al., 2010).

On average, TFA teachers enter teaching following more rigorous academic training and demonstrate higher levels of academic preparation, in particular in subject-specific certification exams, than teachers certified through other pathways. TFA teachers are more likely to come from more elite universities and have taken more rigorous undergraduate coursework than non-TFA teachers (Boyd et al., 2006; Decker et al., 2004a), and outperform teachers trained through other pathways on certification exams testing general knowledge, and liberal arts and science concepts (Boyd et al., 2006). Specific comparisons of math qualifications and outcomes of teachers trained through TFA versus other alternative and traditional pathways suggest that TFA teachers have stronger math credentials, including test scores, coursework, and competitiveness of degree granting institution (Boyd et al., 2010). In high school, TFA teachers also have higher scores on their general content and math-specific certification tests (Xu et al., 2011). Together, this suggests that TFA teachers might be especially effective in high school, particularly in math and science because of their relative academic strengths.

Student baseline proficiency-program fit. In extending the idea of the "fit" between the TFA program and the students it serves, it is important to consider not just the match between the TFA program and the developmental timing of the student, but also the match between the student's readiness and the capacities of the TFA teachers. Work in social-ecological systems theory stresses the importance of the match between the program offerings and the recipients' needs for program success (Cumming, Cumming, & Redman, 2006) and work in early childhood makes similar arguments about the strengths and weaknesses of a variety of child and family interventions (Bradley, McKelvey, & Whiteside, Mansell, 2011; McGuigan, Katzev, & Pratt, 2003).

Bioecological theory also contributes to the conceptualization of the fit between the student and the TFA program. Three core portions of Bronfenbrenner's Process-Person-Context-Time (PPCT) model (Bronfenrenner and Morris 2006, Bron 2005) are particularly relevant when thinking about variation in the effects of TFA based on student-program fit. Here the proximal processes occurring in the context of a TFA classroom, as opposed to counterfactual classrooms, also interact with the qualities of the individual student, including their academic proficiencies.[17] In the case of TFA, the proximal processes that are at the center of the PPCT model, might lead different types of children to respond differently to the same types of program environments created by TFA teachers, and the way that TFA's impacts might vary center around the fit between the program and the child (c.f. E. B. Miller, Farkas, Vandell, & Duncan, 2014). Pianta & Walsh's (1996) Contextual Systems Model (CSM) also highlights the ways in which the fit between the child, their family, the school context, and the culture interact to predict student achievement. A program like TFA, which brings a unique culture of teaching to an existing school environment may have differing impacts based on child factors, including prior

achievement. While I am unable to examine the micro-level interactions and proximal processes that are the focus of the PPCT and the CSM, I am able to examine whether the interaction between the fit of the program in which these processes are occurring varies based on an important child factor, prior achievement.

Several competing hypotheses predict which types of students are most likely to benefit from any type of intervention, that are informative to consider in the case of TFA. The *compensatory* hypothesis (Rutter, 1987; Sameroff & Chandler, 1975) predicts that students who are beginning with the lowest baseline proficiency scores should benefit the most from being in a TFA classroom. These students derive the largest benefits from being in the intensive skill-building environments that TFA classrooms purport to provide. In contrast, the *accumulated advantages* or *skill-begets-skill* (Cunha & Heckman, 2007) paradigms predict that more advantaged children with higher entering proficiency levels will be best positioned to take advantage of the TFA classroom environment and will build upon their existing skills at a faster rate than their peers entering with lower skill levels. Finally, recent work by Miller et al. (2014) finds evidence of a *goldilocks* scenario, in which students who are not too low nor too high in baseline achievement could benefit the most from an intervention like TFA.

Work on the effects of early childhood settings of student academic and socio-emotional development finds evidence of variation in effects based on different levels of student risk factors, finding evidence of the *compensatory* and protective effects of a variety of early interventions. Watamura and colleagues (2011) find that childcare quality is protective against the risk factor of lower-quality home environments, while lower-quality home environments coupled with lower-quality child care environments create a doubly risky setting for children. Similarly, Miller and colleagues (2014) find variation in the effects of Head Start in two

academic domains based on levels of pre-academic stimulation in the home. While in math, Head Start had a *compensatory* effect, with largest effects for students from homes with low parental pre-academic stimulation, in literacy, the largest benefits were for students receiving moderate levels of pre-academic stimulation (the so called *goldilocks* scenario). While both of these studies focus on moderation in early childhood settings, work examining classroom quality in elementary school finds that higher-quality teacher-student relationships buffer children from home-related risk-factors and can help to foster children's achievement, and particularly for initially lower-achieving children (Birch & Ladd, 1997; O'Connor & McCartney, 2007; Pianta, Steinberg, & Rollins, 1995; Pianta, 1999).

In contrast, evidence from other examinations of teacher effectiveness finds confirmation of the *accumulated advantages* hypotheses where teacher effects are both additive and cumulative, but not *compensatory* (Sanders, Wright, & Horn, 1997). Work from early childhood also finds evidence of *accumulated advantages* in the Sure Start program in the United Kingdom, which was more effective for children from more advantaged backgrounds (Belsky et al., 2006).

Finally, some prior evidence suggests that program-student proficiency fit is especially important and can be a source of impact variation. Connor et al. (2006) found evidence of variation in the effectiveness of teaching styles depending on the entering skill level of the students. Teacher-managed reading strategies had a *compensatory* effect for initially-poor readers, while child-managed strategies led to more growth for students with higher initial reading comprehension (i.e. *accumulated advantages*). This suggests that particular program impacts may vary based on the initial skill level of the student, and that different types of interventions may be better suited for students with higher or lower ability levels.

Analytic Approach

To evaluate the stage-proficiency-environment fit of TFA and its effects on student achievement, this chapter examines two research questions:

(1) Does the effect of TFA vary across developmental stages (as operationalized by school levels) and subject domains?

(2) Is this effect moderated by student baseline proficiency?

I examine both research questions by comparing end of year test score outcomes for students in TFA classrooms relative to their peers who had the opportunity to be in a TFA classroom because they attended the same school in the same year, and were in the same grade or took the same course from a non-TFA teacher. I supplement this approach by comparing students to themselves in years (or across subjects within a year) in which they did and did not have a TFA teacher. This is described in further detail below.

To examine whether TFA has differential effects based on baseline proficiency, I follow an approach used by Watamura et al. (2011) and elsewhere in the literature (e.g., E. B. Miller et al., 2014), in which I create baseline test score terciles to test whether TFA's effects on student achievement vary depending on the entering proficiency level of the students. In doing so, I examine the match between the student, their relative level of preparedness, the developmental stage in which they have a TFA teacher, and the context of a TFA classroom versus a non-TFA classroom.

Data

I examine the relationship between TFA, developmental timing, baseline student proficiency and student achievement using administrative data from the state of North Carolina. TFA places teachers in two regions in North Carolina. The Eastern North Carolina region, 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 (teachforamerica.org, 2013b).

Administrative data from North Carolina were provided by the North Carolina Education Research Data Center (NCERDC) at Duke University. NCERDC provided demographic and assessment records for students which were matched to their classroom teachers.[18] These data are well suited to the examination of the effectiveness of TFA versus non-TFA teachers. In addition to the prior studies that directly compare TFA teachers with non-TFA teachers (Henry et al., 2010; Xu et al., 2011), these data have also been used for numerous studies examining teacher quality (Clotfelter et al., 2006), the feasibility and stability of teacher value-added models (Clotfelter, Ladd, & Vigdor, 2007b; Clotfelter et al., 2010; Goldhaber & Hansen, 2010; Rothstein, 2009, 2010), and teacher persistence and turnover (Goldhaber, Gross, & Player, 2011).

All North Carolina TFA teachers (also called corps members) from the entering corps of 1999 to 2010[19] 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).[20] 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, and 679 are included in the focal years of this study. [21]. 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 corps members placed in Eastern North Carolina each year more than doubled.

In Chapter 4 in this dissertation, I examine whether the effects of TFA changed as the TFA program in North Carolina matured. For this analysis, to examine whether TFA's effects change based on the match between the student's developmental timing and academic proficiency, I focus exclusively on the post-2005 period in North Carolina where the TFA program is already thriving. Doing so allows me to focus on TFA's effects once the program is established across subject areas and grade levels so that differences cannot be attributed to TFA's struggles to get a particular subject area of its program under way. TFA corps members by placement region and year for grades K-12 in 2006-2011 are shown in Table 3.1.

[Insert Table 3.1 Here]

From 2006-2011 TFA teachers taught in 431 schools in 73 LEAs and charter schools in North Carolina, with 301 in Eastern North Carolina and 378 in Charlotte, between 2005/2006 and 2010/2011. In the study, I concentrate on the TFA and non-TFA teachers in tested subjects in the same schools, grades, and years. This allows me to directly compare the effects of having a TFA teacher relative to other teachers the same students could have had for the same courses. I describe this comparison in more detail in the methods section below, but to provide more detail about how students and teachers in these two groups differ, a comparison of the demographic characteristics of the students with a TFA teacher versus four different comparison groups is shown in Table 3.2. 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 case, because these students had the potential to be assigned to a TFA teacher in the same year, but were not, and thus it is used to identify counterfactual students and teachers in the analyses described below. The analytic sample throughout this study consists of only those students with TFA teachers in group 1, where there was the potential of being assigned to a TFA teacher in group 2.

[Insert Table 3.2 here]

As indicated in Table 3.2, students of TFA teachers are most similar to other students in non-TFA classrooms in the same schools. They have fewer parents with college degrees or above, more of them are African American, fewer of them were ever identified, and more of them are over-age for grade or have repeated a grade than students in other schools and LEAs. They also have more students from the bottom third of the baseline test score distribution than teachers in the other groups.

The outcome of interest in this study is end-of-grade or end-of-course student test scores. In grades three through eight, tested subjects include math and reading, which I combine across grades 3-5 and 6-8 for parsimony.[22] In high school, tested subjects include math, English, science, and social studies. Three of these four global subjects consist of several individual tested subjects, all of which are combined for 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 consist of U.S. History and Civics and Economics; and English consists of English 1. Not all of the individual subjects are tested each year, but at least one subject in each global subject area was tested every year (North Carolina Department of Public Schools, 2011).

Student test scores are standardized within year and grade for end-of-grade tests and within year and subject for high school end-of-course tests. 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. Using the lagged standardized scores, I create test score terciles for students with non-missing lagged scores to represent students' relative proficiency levels as they enter TFA or non-TFA classrooms. For high school outcomes, as students do not take a consistent course sequence in the same grades across the state, I create terciles of eighth grade math and reading scores.

All models include the student controls listed in Table 3.2. In an effort to minimize missing data, the student characteristics were filled in from across several individual datasets provided by the NCERDC. For student gender and race ethnicity, the modal value for an individual across these datasets was taken and filled in cases that it was missing. For example, if student gender was present in some years and missing others, or if it agreed in three out of four total years the student was in the dataset, the modal non-missing value was taken for all four years.

Using the data filled in across several data sources, I created the 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 taken. Parental education 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 they were never identified as having any of these three characteristics. 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. The student the over-and under-age variables are not fixed from year to year and can change following a retention or early promotion, unlike the rest of the student characteristics which are filled in across multiple years of student observations.

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 quality 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 whether TFA's effects are moderated by the developmental timing and baseline student proficiency level, I use two empirical approaches. The first approach uses fixed effects for school, grade, and year combinations and the second uses student fixed effects.[23] 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 (or school, grade, year in high school, but for parsimony going forward, I list school, grade, and year to refer to both sets of fixed effects). 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 primary evaluation of achievement gains in TFA versus non-TFA classrooms will be 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 (c.f. Table 3.2).

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:

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

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 linear function of the student's prior-year (or eighth grade) test score, 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. Models testing whether the effect of TFA varies by student incoming skill level include terciles for baseline test scores in place of the

linear terms for prior year or eighth grade test scores. These terciles are created from lagged or eighth grade test scores with the middle tercile omitted as the reference category. These models also include an interaction between baseline tercile and TFA to test whether TFA's effects vary across the baseline test score distribution.

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 grouping makes comparisons within school-grade-year units so 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 eliminates 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 or subjects 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 grade level that may be due to unobservables related to the student, the teacher, or parental preferences. The presence and magnitude of this bias in North Carolina and other settings continues to be debated (Clotfelter et al., 2006; Goldhaber & Hansen, 2010; Koedel & Betts, 2011; Rothstein, 2009), however this approach continues to be used in many administrative data applications for comparing the impact of TFA teachers on student achievement (e.g., Boyd et al., 2009, 2006).

To address some of these biases, I also use a second approach to identify the relationship between TFA and student achievement. In addition to comparing students in the same school, grade, and year that do and do not have TFA teachers, the second analytic approach uses students as their own counterfactuals. To do so, I use student fixed effects among the group of students that ever have a TFA teacher, comparing their outcomes in the years in which they do and do not have a TFA teacher. This is similar to the approach used by Clotfelter, Ladd, and Vigdor (2007a) who use student fixed effects with the North Carolina data to examine whether teacher characteristics (but not whether teachers were TFA) impact student achievement in elementary school. I thus estimate the following model:

$$A_{iy} = \gamma_o + \gamma_1 TFA_{iy} + \pi_i + \varepsilon_{iy} (2)$$

where A_{iy} is an individual student i's standardized achievement level in year y, and is a function of whether or not they have a TFA teacher in that year, TFA_{iy}, plus a student specific fixed effect, π_i. I estimate two versions of these models, one in which I compare students with themselves in any year in which they have a non-TFA teacher and a second version in which I compare them only relative to years prior to the first time they have a non-TFA teacher. I also estimate a similar version of this model to compare TFA effects within a given year across high school subjects following the approach used by Xu et al. (2011) in a previous evaluation of TFA's effects in high school in NC. This analysis is largely a replication of the previous analysis by Xu and colleagues, but it does represent an important part of testing my research questions.

These student fixed effects approaches address some of the concerns about bias that cannot be resolved by the school-grade-year fixed effects models. Specifically, they eliminate bias due to non-random sorting of students into classrooms that is based on time-invariant characteristics of the student. The student fixed effects models further control for time varying

characteristics within each student. However, it is important to note that the student fixed effects are not a panacea, and there may still be some bias from time varying student characteristics that result in the student being placed in a TFA classroom in a given year or in a TFA classroom for one subject and not another.

Results

To evaluate whether the effects of TFA vary across developmental stages and baseline proficiency, I first examine the relationship between TFA and student achievement at different developmental stages and in different subjects using school-grade-year (or school-subject-year) fixed effects. As described in the methods section, these models compare students who have a TFA teacher with students who have a non-TFA teacher in the same school, grade level (or subject area), and school year. The results of these models are shown in Table 3.3.

[Insert Table 3.3 here]

Table 3.3 shows that the effects of TFA are positive in most subjects and across most grade levels, although they differ in magnitude. TFA effects are by far the largest in high school science, where having a TFA teacher improves student achievement by 0.197 standard deviations. The TFA effect in math is also fairly substantial at 0.112 standard deviations. Effects in high school English, social studies, and elementary and middle school math are all similarly sized, ranging from 0.036 standard deviations to 0.050 standard deviations. The one exception to these positive effects is elementary reading, where the TFA coefficient is negative and marginally significant.

Additional models testing whether these broad patterns hold for individual subjects and grades are shown in Appendix Tables 3.1 and 3.2. These results show that across all elementary and middle school grades, TFA has a positive effect in math. In contrast, the results across

grades are inconsistent for elementary and middle school reading, where the TFA effect is positive in 8th grade, negative in 5th grade, and indistinguishable from zero in the remaining three grades. Looking across subjects in high school, TFA has positive effects in every high school math subject (with the largest effects (0.163 SD) in geometry), and in three of the four science subjects (with the largest effects (0.355 SD) in chemistry). Results are more mixed in social studies, with large negative effects in Economics, Legal, and Political Systems[24] (-0.971 SD), and moderate positive effects in the other two subjects. There is only one English test in high school, rendering further examination unnecessary. Thus, overall, the results from the individual subject and grade level models generally conform to the broader results presented in Table 3.3, though there are some notable exceptions.

Table 3.4 shows results from regressions that formally test for differences in TFA's effects in math and reading across school levels by including interaction terms for TFA and middle and high school. The results in Table 3.4 show interesting differences across math and reading. The main effect of TFA in math shows that elementary school students with TFA teachers score 0.054 standard deviations higher in math than elementary school students with non-TFA teachers. The non-significant, near-zero coefficient for middle school suggests that the effects of TFA in elementary and middle school differ very little. However, TFA's effects are 0.047 standard deviations higher in high school, suggesting that in math, TFA has a positive effect in general, but has a particularly large positive effect in high school.

In contrast, in English language arts (ELA), TFA does not have a significant effect in elementary school, and the point estimate for TFA is nearly zero. The difference between the effects in elementary and high school is also small and not statistically significant (as shown by the TFA*high school coefficient). However, TFA effects in ELA do appear to differ between

elementary and middle school, with middle school students gaining .037 standard deviations more from having a TFA teacher than elementary school TFA students, suggesting that there is some benefit to having a TFA teacher in middle school. These findings suggest that the ideal time to have a TFA teacher may differ depending on whether one is primarily concerned with boosting math or ELA scores, as TFA's effectiveness across subjects differs across developmental time periods.

[Insert Table 3.4 here]

While Table 3.4 uses school-grade-year (school-subject-year in high school) fixed effects to control for all shared characteristics of being in a particular school, in a particular grade (or subject) in a particular year, Table 3.5 instead uses student fixed effects to account for all characteristics of an individual that are constant across time. In these models, instead of comparing across same-grade students within a school and year, I estimate the effects of TFA comparing students with themselves in grades in which they did and did not have a TFA teacher. One of the benefits of using a student fixed effects approach is that these results take into account a wider range of individual characteristics including unobserved characteristics that are consistent within a student across time.

Estimates might differ across these two types of models if there are unobserved student characteristics that lead some students to be sorted into a TFA classroom. For example, one can imagine veteran teachers arranging class assignments so that all of the best-behaved students are in their classes, leaving the students with problem behaviors for a novice (TFA) teacher. This might lead to bias in the estimates provided by the school-grade-year fixed effects due to behavioral characteristics that are not observable in these data. However, the results reported in Table 3.5 compare students assigned to TFA teachers to themselves in other school years, so that

issues like chronic problem behaviors are controlled for in the estimates. While there are other characteristics, such as temporary changes in problem behavior due to a tragic life event, which cannot be accounted for in the student fixed effects models, such models are able to control for more unobservable characteristics than the school-grade-year fixed effects.

The student fixed effects results reported in Table 3.5 provide a somewhat different picture of the pattern of variation in the effects of TFA than the results of the school-grade-year fixed effects models reported in Table 3.4. First, we see that in both math and ELA, the average effects of having a TFA teacher do not differ between elementary and middle school, though the point estimates suggest that TFA teachers are perhaps less efficacious in middle school math and also less problematic in middle school reading. Second, we see that the effects of having a TFA teacher in both Math and ELA are significantly larger in high school than in elementary school. Third, and perhaps most interestingly, the estimate of the effect of TFA in elementary school reading is now negative and significant, so that when we compare individuals in years when they have TFA and non-TFA teachers we see that they do worse in reading in the years that they have TFA teachers. (Note that while the point estimate in Table 3.4 is quite close to zero, the point estimate in Table 3.3 is negative and marginally significant.) Taken together, the results presented in Tables 3.3-3.5 suggest significant variation in the relative effectiveness of TFA across developmental stages, with high school students benefitting the most from having a TFA teacher, while elementary and middle school students benefitting about equally (though in reading middle school students appear to benefit more from TFA teachers, or perhaps be hurt less by them, than in elementary school).

[Insert Table 3.5 here]

Table 3.6 shows results for high school only, using a student-year fixed effects approach to compare TFA's effects across subjects within the same year. These models compare students with themselves in the same year across different high school subjects, taking advantage of the fact that even within a given year students have TFA teachers in some subjects and not others. The results comparing TFA effects across high school subjects largely confirm the school-subject-year fixed effect findings from Table 3.3. TFA has a small positive effect in high school English, which is somewhat sensitive to model specification. The TFA effects in the other three subjects are larger than those in English, with TFA science teachers having the largest effects on student achievement. These models suggest that, even within a developmental stage, TFA's effects on student outcomes vary depending on the subject matter.

[Insert Table 3.6 here]

To examine the second research question, which focuses on the fit between student proficiency and TFA, the models reported in Table 3.7 replicate the approach of Table 3.3 (using school-grade-year, or school-subject-year fixed effects), and examine the effects of TFA across different subjects and developmental stages. However, the results of Table 3.7 build on those in Table 3.3 by including interactions between having a TFA teacher and being in the bottom or top tercile of baseline test scores (the middle tercile is the reference category). These interaction effects test whether TFA has differential effects based on students' entering proficiency level.[25]

[Insert Table 3.7 here]

The results in Table 3.7 show a remarkably consistent and countervailing pattern between elementary and middle school students on the one hand and high school students on the other. In elementary and middle school, across both math and reading, TFA's effects conform to the *accumulated advantages* hypotheses, with high proficiency students benefitting the most from

having a TFA teacher. In math, TFA has positive average effects in both elementary and middle school that do not differ for students with entering skill levels that are in the bottom two thirds of the baseline test score distribution. However, TFA has even larger positive effects for students who enter with test scores in the highest third of the distribution, so that elementary school students who scored in the top tercile in the prior year gain an additional 0.082 standard deviations in math from having a TFA teacher, while middle school students who scored in the top tercile gain an additional 0.051 standard deviations. In reading, the main effects of TFA is not different than zero in the bottom two terciles of the baseline test score distribution (although these point estimates are close to the margin of significance and the signs are in opposite directions, consistent with other models showing a marginally significant negative effect in elementary school and a significant positive effect in middle school). However, the effects for students who start in the top third of the baseline reading distribution are positive and even larger than the *accumulated advantage* effects in math (with coefficients of 0.132 in elementary reading and 0.090 in middle school reading). These findings echo those of Chapter 2, where TFA teachers had a negative impact on the bottom of the post-test distribution. Both sets of result suggest that TFA teachers' struggles with reading instruction are concentrated among low achievers, while high achieving readers may benefit from being in a TFA classroom.

While the elementary and middle school results suggest an *accumulated advantages* perspective, the results for high school are instead consistent with the *compensatory* hypothesis, as high school TFA teachers have the largest effects for students who enter in the bottom of the test score distribution in all four high school subjects. In high school math and science, the main effects of TFA are fairly large and do not differ across the top two baseline test score terciles, but students from the bottom tercile gain an additional 0.051 standard deviations from having a TFA

math teacher and 0.064 standard deviations from a TFA science teacher. Interestingly, while students in TFA English and social studies classes in the top two terciles perform no differently than their peers in non-TFA classrooms, students from the bottom third of the baseline test score distribution benefit 0.045 standard deviations more in English and 0.081standard deviations more in social studies.[26]

Finally, to test the stage-proficiency-environment-fit, which examines variation in the effectiveness of TFA across both developmental stages and baseline proficiency levels, I estimate models interacting TFA, school level, and baseline test score tercile for math and ELA. The full results are reported in Appendix Table 3.3 and are summarized in Figures 3.1 and 3.2.

[Insert Figures 3.1 and 3.2 here]

Figure 3.1 contrasts the effects of TFA on math achievement in elementary, middle, and high school across students with baseline test score from different terciles. The figure echoes the results documented in Table 3.7, showing that the pattern for TFA effectiveness across baseline proficiency operates differently in high school than it does in elementary school and middle school. In elementary and middle school, TFA has small positive effects of about 0.06 standard deviations for students in the bottom two baseline test score terciles. But for the upper tercile, TFA has a larger effect of roughly one tenth of a standard deviation. In other words, in elementary and middle school math, TFA has some positive impact on all students, but it follows the *accumulated advantages* hypothesis, showing the largest effects for the children with the highest entering skill levels. In high school, TFA also has positive effects across the baseline test score distribution. However, in contrast to elementary and middle school, where TFA's effects were largest for the best-prepared students, in high school the opposite is true. TFA's effects on students in the top two terciles are just under one tenth of a standard deviation, where in the

lowest tercile, they are above 0.13 standard deviations. This suggests that TFA's effects in high school math follow the predictions of the *compensatory* hypothesis.

Figure 3.2 summarizes the results of the same regression model predicting language arts, showing the same general pattern of the math results, but with a slightly different intercepts. As with math, the elementary and middle school TFA effects conform to the same pattern where TFA's effects are largest for students entering in the highest tercile, which are near the effects found for top tercile math students, at roughly 0.08 standard deviations in elementary school and 0.12 standard deviations in middle school. However, TFA has small positive ELA effects in middle school (0.02 standard deviations) for students in the bottom two entering terciles and small negative ELA effects for middle and lower tercile elementary school students. High school ELA also follows a similar pattern to math, which has positive effects across the baseline test score distribution, but with the largest effects among students from the bottom tercile. However, the ELA effects are smaller than the math effects, and are less than 0.08 standard deviations even among the bottom tercile students. Thus, the ELA results also suggest that in elementary and middle school TFA follows an *accumulated advantages* hypothesis, while in high school it has a *compensatory* effect on student achievement.

Discussion

This chapter examines whether TFA's effects on student achievement vary across developmental stages, subject areas, and by students' entering skill levels using administrative data from the state of North Carolina. The results suggest that the impact of TFA is heterogeneous and varies across all three of these dimensions. First, looking across subjects I find that TFA has its largest effects in mathematics and science, with smaller (sometimes null) effects in reading and social studies. Looking next at different developmental stages, reveals that

in general, students benefit the most from having a TFA teacher in high school. Not surprisingly, then, I find that the largest effects of having a TFA teacher are in high school math and science, a finding that is congruent with previous research in North Carolina and elsewhere (Glazerman et al., 2006; Henry et al., 2010). Likewise, the one area in which TFA has null or potentially negative effects is in elementary reading, a finding that is consistent with my results from Chapter 2.

The third dimension of heterogeneity in the effectiveness of TFA teachers that I examine, with perhaps the most interesting results, is the students' baseline proficiency. These results highlight the importance of considering the person-environment fit of an intervention, as they show that the relationship between having a TFA teacher and a student's baseline proficiency varies across developmental stages. In elementary and middle school, TFA's effects are largest for students with the highest entering proficiency levels, following the *accumulated advantages* or *skill-begets-skill* hypotheses. In math TFA has small positive effects, even for students with the lowest entering skill levels, while for reading, the effects of TFA are, if anything, slightly negative. This is congruent with my findings from Chapter 2, and suggests that TFA struggles to reach students most in need of literacy intervention in elementary school. This particular finding also underscores the need to examine TFA's training related to early literacy and literacy intervention and point to an area in which teacher experience and traditional pathways to teacher certification are better than the TFA program.

In contrast, in high school, TFA's effects are largest for students with the lowest entering skills, and it thus appears to follow the *compensatory* hypothesis. While TFA effects are positive overall in high school, and are especially strong in math and science, they are even higher for the initially less-advantaged students. This suggests that TFA training, TFA teachers' content

knowledge, or their intervention methods helps to support struggling students, at least relative to the teacher training and intervention methods of counterfactual teachers. Future research would do well to examine TFA classrooms and training programs to investigate the ways in which TFA teachers foster this type of support for their initially low-performing students. While it is ultimately unclear which is preferable, given TFA's emphasis on equity, an argument can be made that a *compensatory* approach is preferable, in which case TFA teachers in high school are potentially doing a "better" job. However, given that the vast majority of TFA students are disadvantaged, helping the higher performing students presumably also contributes to the equity goals of TFA, so that the most desirable pattern of results from the program's perspective might be one in which all students benefit equally. Such speculation about the ideal pattern of results, however, is well beyond the scope of this study and is best left to normative rather than empirical analysis. However, by deepening our understanding of where TFA works best these results provide insight into how TFA works (or where traditionally trained teachers struggle), so that we can potentially design better interventions and teacher training programs that will help meet the needs of all students.

These findings conform to some of the hypothesized patterns of results, but not others. The literature on emotionally supportive school structures and curricular integration predicted that the largest effects would be found in elementary school. Instead, TFA effects were weakest, and in some cases negative, in elementary school, suggesting that these factors do not differentially support TFA teachers. Drawing from the stage-environment fit literature, I predicted that TFA's effects would be largest in middle school. Instead, they were larger in high school. Inferring from the literature on teacher content knowledge, I predicted that TFA teachers would perform most favorably in later grades where content knowledge becomes more

specialized, and in technical subjects, such as math and science. This hypothesis is supported by my results. The findings further suggest that TFA teachers' relative comfort with the technical nature of high school math and science might outweigh the relative benefits of having creating a supportive environment or integrated curriculum, leading them to have larger effects in high school than in middle school. While this comparison of mechanisms is not possible here, it does seem that TFA's relative strength in math and science contributes to greater effectiveness in these subjects across grade levels, which is in keeping with previous research across subject areas (Henry et al., 2010; Xu et al., 2011).

The literature was less clear about the predicted pattern of results with regard to baseline proficiency. There was evidence from literatures on educational interventions and teacher quality that had conflicting findings about whether to expect larger effects for initially lower-, middle-, or higher-performing students. Thus, perhaps it is not surprising that there was a mixed pattern of effects in terms of relative effectiveness for baseline proficiency levels with divisions across school levels.

These results also point to the need for further research about the lasting effects of the observed relationships identified here. The short term effects may not have lasting impacts on students' achievement or attainment, and these effects may also vary across the dimensions of developmental stage and student proficiency. As the evidence about the persistence of intervention and teacher effects and the debate about the ideal timing of intensive interventions persists (Barnett, 1996, 2011; Chetty, Friedman, Hilger, et al., 2011; Jacob, Lefgren, & Sims, 2010; Knudsen, Heckman, Cameron, & Shonkoff, 2006; Knudsen, 2004; Lee, Brooks-Gunn, Schnur, & Liaw, 1990; Lee & Loeb, 1995; L. B. Miller & Bizzell, 1983), this will be an important avenue for future work examining TFA effectiveness.

This chapter also suffers from several limitations. As with all observational data, the conclusions generated here, while generalizable to the population of students in TFA schools in North Carolina, do not generalize more broadly. They effects are also not causal estimates of TFA's effects as students were not randomly assigned to teachers. However, the results do represent the best attempt to understand the types of impacts TFA teachers have on student achievement in a real-world setting, across an entire state for several years. The types of selection into schools and classrooms that are present in these data are akin to the same types of issues that any intervention would face when scaled up into real world schools in which investigators do not have complete control. The controls and fixed effects included in the program estimates have helped to reduce bias stemming from selection and have also helped to restrict the comparison of TFA teachers and their students to counterfactual teachers and students in similar settings, rather than across all types of schools and districts, to which TFA students do not have access. Likewise, it is reassuring that models with student fixed effects, which account for all unmeasured characteristics of an individual that remain constant over time (or across subject), yield similar findings. Thus, while not providing a strictly causal estimate, these results represent a rigorous attempt to identify TFA's effects within a real state-wide context.

In addition, one important concern in the middle and high school analyses is that the lowest-performing students have already dropped out of school and thus TFA's *compensatory* effects are only for the moderately disadvantaged rather than the very worst off. This is a valid concern that is difficult to deal with, however, this study again represents how the TFA intervention operates in a full-scale context that includes students who have already removed themselves from school. That is, as it is difficult to think about how a school-based intervention could affect students who do not attend, such interventions and policies targeted at different

school levels are by definition targeting different students, and evaluations of their relative efficacy is perhaps best conceptualized on the basis of the students they are targeting.

In sum, these results suggest that TFA's effects, while varied, are largely positive. Although the TFA program continues to have its critics, administrators in North Carolina in low-performing schools should not rule out the TFA program when seeking new teachers. However, they should consider ways to support TFA teachers working with lower-performing readers. Further work is needed to investigate the aspects of the TFA training program and teacher practices that yield these positive effects, and further work should seek to investigate mechanisms leading to variation in program impacts, the beginnings of which are examined in Chapter 4. Finally, this research underscores the need for future work examining TFA's effects should continue to consider the ways in which program impacts vary based on the stage-proficiency-environment-fit between TFA and the students it serves.

---

[16] TFA teachers are recruited to a district by TFA through a partnership contract for a given number of teachers assigned to a particular subject. Partner districts typically pay TFA a fee for each corps member recruited, particularly if the corps member will be staffing a high-need subject area. This fee varies across different sites, but ranges from $2,000 to $5,000 per teacher (Simon, 2012).

[17] The timing aspect of a student's experience in a TFA classroom is also important, although in this theoretical framework, Bronfenbrenner focuses more about the chronological and historical timing rather than the developmental timing.

[18] Technically, these 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. Clotfelter, Ladd, and Vigdor (Clotfelter, Ladd, & Vigdor, 2007a) examine the accuracy of these matches for elementary school in detail, also

examining whether the teacher matched a personnel record and taught a valid math or reading course code in the relevant subject in the same year.

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. For parsimony, I currently refer to the test administrator as the teacher.

[19] 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.

[20] 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 2012/13 there were 115 LEAs and 107 charter schools (http://www.ncpublicschools.org/docs/fbs/resources/leacharterlist.pdf, 2013).

[21] 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 4.2.

[22] 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 grade 6-8 in my analyses.

[23] A third approach, which uses teacher fixed effects, 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).

[24] It is important to note that the exams in Social Studies have 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. The remaining exam subjects have been mostly constant over time, but were updated in either 2005/2006 or 2006/2007 to increase their difficulty (North Carolina Department of Public Schools, 2011).

[25] I do not employ student or student-year fixed effects for models examining the accumulated advantages and compensatory hypotheses, as these hypotheses predict that students' gains will vary as a function of why they stand relative to their peers, and it is not clear that we would expect that being high or low relative to other observations within the person (or person-year) would function similarly. Thus, for example, if an individual had three observations, and scored at the 97th, 98th, and 99th percentiles, (or 1st, 2nd, and 3rd percentiles) it is not clear that we should treat the first as being a bottom tercile observation and the last as a top tercile observation, as it is not theoretically clear that the compensatory and accumulated advantages frameworks would predict differential growth on the basis of these within-student differences.

[26] Results (not shown) testing the same relationship with a median split confirm these relationships.

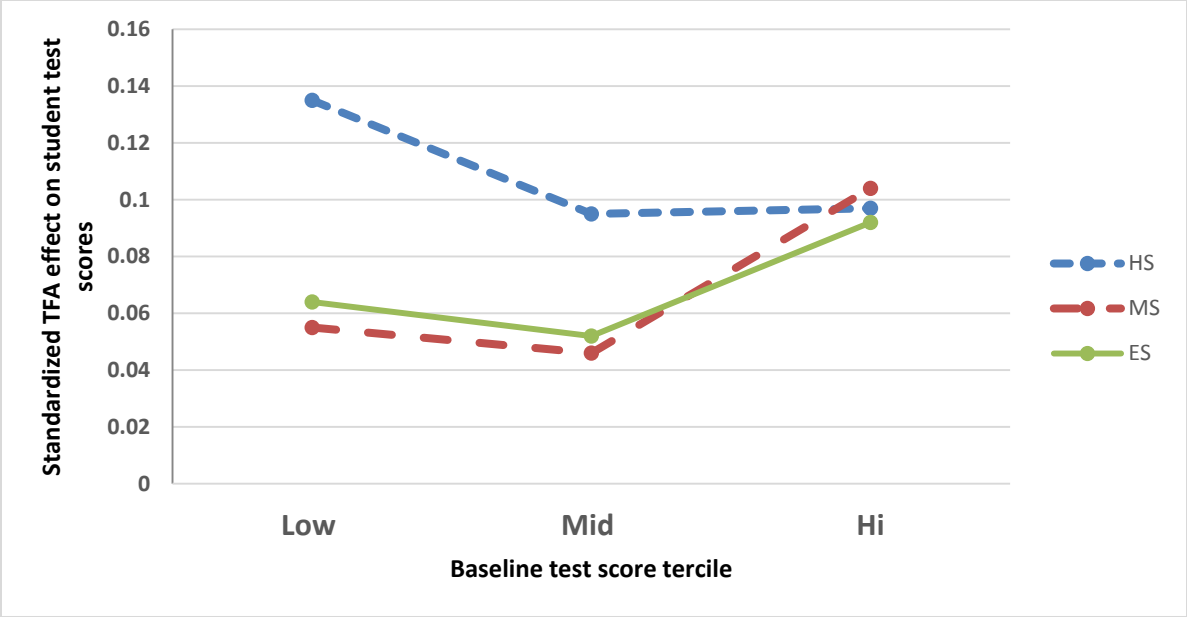


Figure 3.1. TFA effects on math achievement across developmental stages, by baseline proficiency levels.

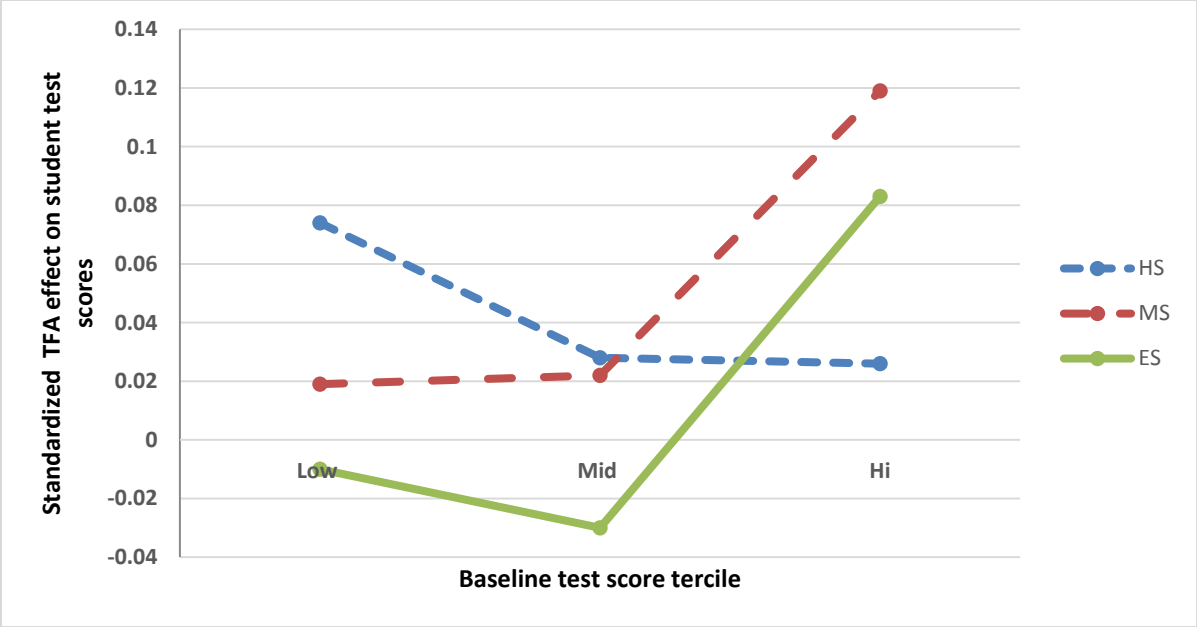


Figure 3.2. TFA effects on reading achievement across developmental stages, by baseline proficiency levels.

Table 3.1.

Number of TFA teachers placed in tested subject classrooms by region and year, grades K-12, 2006-2011

Year	Charlotte	Eastern North Carolina	Total
2006	58	60	118
2007	55	51	106
2008	98	82	180
2009	63	56	119
2010	104	52	156
Total	378	301	679

Table 3.2.

Student and teacher characteristics across comparison samples, 2006-2011

	TFA Students	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 Gender						
Male	50.1	50.2	50.4	50.5	50.6	50.0
Female	49.9	49.8	49.6	49.5	49.4	49.1
Missing Gender	0.0	0.0	0.0	0.0	0.0	0.9
Parent Education Level						
High school or less	24.8	20.1	16.3	16.6	22.0	17.9
Some college/trade school	23.6	20.4	17.5	17.9	22.8	19.0
College degree or higher	21.9	34.5	39.8	33.5	27.7	32.9
Missing Parent Ed.	29.7	25.0	26.4	32.0	27.6	30.2
Ever identified as LEP						
Never identified	86.3	86.2	88.5	91.2	90.6	89.4
Ever identified	11.7	11.7	9.8	7.6	8.4	8.4
Missing LEP information	1.9	2.1	1.7	1.2	1.0	2.2
Student Ethnicity						
White	14.1	28.3	45.3	56.4	69.4	55.1
Black	67.8	52.8	37.3	27.0	16.7	27.9
Hispanic	11.4	11.5	9.9	8.4	8.5	8.8
Asian	2.5	3.2	3.4	2.5	1.5	2.4
American Indian	0.8	0.6	0.6	2.0	1.1	1.4
Other ethnicity	3.2	3.6	3.5	3.8	2.8	3.4
Ethnicity missing	0.0	0.0	0.0	0.0	0.0	0.9
Has some type of disability						
No	82.6	82.5	83.0	82.5	81.9	81.7
Yes	15.6	15.6	15.5	16.5	17.2	16.3
Missing disability status	1.8	1.9	1.5	1.0	0.8	2.0
Ever identified as gifted						
No	88.3	81.7	77.0	76.5	80.2	77.2
Yes	9.9	16.3	21.4	22.3	18.9	20.7
Missing gifted status	1.9	2.0	1.6	1.2	1.0	2.2
Over-age for grade						
No	68.4	69.3	75.2	76.4	76.5	75.0
Yes	31.6	29.3	23.7	22.7	22.6	23.1
Missing birthdate	0.0	1.4	1.1	0.9	0.9	1.9
Under-age for grade						
No	97.9	96.2	96.9	97.6	97.9	96.6
Yes	2.1	2.4	2.0	1.6	1.2	1.6
Missing birthdate	0.0	1.4	1.1	0.9	0.9	1.9
Ever repeated grade						
No	95.9	96.2	97.4	98.1	98.4	97.0
Yes	4.0	3.7	2.5	1.8	1.5	2.0
Missing repeater status	0.1	0.1	0.1	0.1	0.1	1.0

Administrative data from North Carolina provided by the NCERDC; An LEA is a Local Education Authority, akin to a school district.

Table 3.2 continued

Student and teacher characteristics across comparison samples, 2006-2011

	TFA Students	Non-TFA, Same School-Grade-Year	Non-TFA, Same School, Different Grade/Yr	Non-TFA, Same LEA, Different School	Non-TFA, Different LEA	Total
Baseline Math (%)						
First Tercile	45.17	37.91	30.31	27.85	29.21	29.33
Second Tercile	23.68	23.63	23.10	23.31	25.49	23.82
Third Tercile	14.17	20.67	27.94	27.85	27.87	27.35
Missing Baseline Math	16.99	17.79	18.64	20.99	17.43	19.50
Baseline Reading						
First Tercile	46.52	38.13	30.13	27.83	29.60	29.40
Second Tercile	22.56	23.25	23.32	23.32	25.17	23.76
Third Tercile	13.69	20.57	27.67	27.74	27.60	27.17
Missing Baseline Math	17.23	18.04	18.88	21.12	17.64	19.68
N (Students)	104,715	1,057,729	3,031,257	8,913,570	4,152,364	17,259,635
N (Teachers)	776	3,944	14,345	39,788	18,244	77,097

Administrative data from North Carolina provided by the NCERDC; An LEA is a Local Education Authority, akin to a school district.

Table 3.3.

Impact of TFA on achievement in elementary, middle, and high school, with school-grade-year or school-subject-year fixed effects

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.040*****	-0.019	**0.046*****	**0.023*****	**0.036*****	**0.112*****	**0.197*****	**0.050*****
	(0.010)	(0.010)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.253***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.566***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score Missing	-0.343***	0.478***	-0.328***	0.514***	0.361***	-0.107***	0.137***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.198***	-0.186***	-0.283***	-0.280***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.209***	0.102***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.103***	-0.142***	-0.112***	-0.164***	-0.177***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.027***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.125***	0.167***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.154***	0.139***	0.159***	0.173***	0.158***	0.162***	0.193***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.327***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.417***	-0.090**	-0.063	-0.238***
	(.)	(.)	(.)	(.)	(0.032)	(0.030)	(0.033)	(0.029)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 3.3. continued

Impact of TFA on achievement in elementary, middle, and high school, with school-grade-year or school-subject-year fixed effects

	Elementary Math	**Elementary Reading**	**Middle School Math**	**Middle School Reading**	**High School English**	**High School Math**	**High School Science**	**High School Social Studies**
Student is gifted	0.712***	0.616***	0.400***	0.304***	0.326***	0.336***	0.357***	0.370***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.615***	-0.429***	-0.538***	-0.590***	-0.242**	-0.353***	-0.194	-0.287***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.146***	-0.246***	-0.320***	-0.362***	-0.505***	0.080	-0.218**	-0.301***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.170***	-0.168***	-0.129***	-0.121***	-0.202***	-0.189***	-0.237***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.635***	-0.630***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.280*	-0.823***	-0.461***	-0.604***	-0.249***	-0.409***	-0.672***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.075***	0.013***	-0.020***	-0.029***	-0.004*	-0.161***	-0.037***	-0.016***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.533	0.501	0.657	0.605	0.594	0.428	0.415	0.429
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 3.4

Impact of TFA on same-year tests by school level, with school-subject-year fixed effects

	Math	**ELA**
TFA	**0.054*** **	0.005
	(0.012)	(0.012)
TFA*Middle School	0.003	**0.037***
	(0.015)	(0.015)
TFA*High School	**0.047****	0.019
	(0.016)	(0.018)
Black	-0.398***	-0.414***
	(0.001)	(0.001)
Hispanic	-0.094***	-0.139***
	(0.002)	(0.002)
Asian	0.239***	-0.023***
	(0.002)	(0.003)
American Indian	-0.210***	-0.235***
	(0.003)	(0.004)
Other Race	-0.146***	-0.122***
	(0.002)	(0.002)
Female	-0.070***	0.066***
	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.101***	0.139***
	(0.001)	(0.001)
Parent Ed: College or more	0.241***	0.309***
	(0.001)	(0.001)
Parent Ed: Missing	0.132***	0.162***
	(0.002)	(0.002)
Student has disability	-0.372***	-0.492***
	(0.001)	(0.001)
Disability status is missing	0.008	-0.357***
	(0.035)	(0.043)
Student is gifted	0.973***	0.901***
	(0.001)	(0.001)
Gifted status missing	-0.630***	-0.513***
	(0.025)	(0.038)

* p<0.05 ** p<0.01 *** p<0.001; Baseline tests are lagged test scores in grades 4-8 and 8th grade tests in grades 9-12; Omitted categories are: White; Parent Ed. High Sch. or less; Elementary School; Standard errors in parentheses.

Table 3.4 continued

Impact of TFA on same-year tests by school level, with school-subject-year fixed effects

	Math	**ELA**
Student ever IDed as LEP	-0.136***	-0.349***
	(0.002)	(0.002)
LEP status missing	-0.235***	-0.330***
	(0.020)	(0.022)
Student over age for grade	-0.324***	-0.314***
	(0.001)	(0.001)
Student under age for grade	0.107***	0.069***
	(0.003)	(0.004)
Student has repeated grade	-0.798***	-0.747***
	(0.003)	(0.003)
Repeater status missing	-0.718***	-0.822***
	(0.036)	(0.051)
Constant	-0.057***	-0.048***
	(0.001)	(0.001)
R-squared	0.357	0.375
Observations	5261101	4533172

* p<0.05 ** p<0.01 *** p<0.001; Baseline tests are lagged test scores in grades 4-8 and 8th grade tests in grades 9-12; Omitted categories are: White; Parent Ed. High Sch. or less; Elementary School; Standard errors in parentheses.

Table 3.5

Impact of TFA on same-year tests across school levels, with student fixed effects

	Compared with any year		**Compared with pre-1st TFA year only**	
	Math	**English Language Arts**	**Math**	**English Language Arts**
TFA	**0.061*** **	**-0.017***	**0.056*** **	**-0.021****
	(0.008)	(0.007)	(0.009)	(0.007)
Middle School	-0.047***	-0.021***	-0.052***	-0.022***
	(0.006)	(0.005)	(0.006)	(0.005)
High School	-0.188***	0.029***	-0.194***	0.028***
	(0.006)	(0.005)	(0.006)	(0.005)
TFA* Middle School	-0.018	0.016	-0.021	0.015
	(0.011)	(0.009)	(0.011)	(0.009)
TFA* High School	**0.102*** **	**0.084*** **	**0.116*** **	**0.082*** **
	(0.011)	(0.011)	(0.012)	(0.012)
Constant	0.055***	-0.003	0.062***	-0.001
	(0.006)	(0.005)	(0.006)	(0.005)
R-squared	0.019	0.004	0.019	0.004
Observations	2698177	1985583	2691971	1981563

* p<0.05 ** p<0.01 *** p<0.001; Omitted category is: Elementary School.

Table 3.6

Impact of TFA on same-year tests across high school subjects, with student-year fixed effects

	Compared with any year	**Compared with pre-1st TFA year only**
TFA	**0.032****	0.022
	(0.010)	(0.011)
High School Math	-0.158***	-0.158***
	(0.001)	(0.001)
High School Science	-0.023***	-0.023***
	(0.001)	(0.001)
High School Social Studies	0.015***	0.015***
	(0.001)	(0.001)
TFA*HS Math	**0.082*****	**0.074*****
	(0.012)	(0.014)
TFA*HS Science	**0.200*****	**0.144*****
	(0.013)	(0.015)
TFA*HS Social Studies	**0.053*****	**0.065*****
	(0.014)	(0.017)
Constant	0.020***	0.021***
	(0.001)	(0.001)
R-squared	0.026	0.026
Observations	7924846	7913068

* p<0.05 ** p<0.01 *** p<0.001; Omitted category is: High School English.

Table 3.7

Impact of TFA on same-year tests based on baseline proficiency, with school-subject-year fixed effects

	Elementary Math	Elementary Reading	Middle Sch. Math	Middle Sch. Reading	High School Math	High School English	High School Science	High Sch. Soc. Studies
TFA	**0.052****	-0.028	**0.047*****	0.020	**0.092*****	0.027	**0.174*****	0.004
	(0.020)	(0.021)	(0.011)	(0.011)	(0.013)	(0.017)	(0.016)	(0.017)
TFA*Bottom third baseline test scores	0.007	0.013	0.009	0.001	**0.051*****	**0.045***	**0.064*****	**0.081*****
	(0.023)	(0.024)	(0.012)	(0.013)	(0.015)	(0.019)	(0.018)	(0.020)
TFA*Top third baseline test scores	**0.082****	**0.132*****	**0.051****	**0.090*****	-0.030	-0.009	0.012	0.025
	(0.031)	(0.032)	(0.016)	(0.016)	(0.023)	(0.027)	(0.027)	(0.030)
Bottom third baseline test scores	-0.546***	-0.570***	-0.531***	-0.601***	-0.420***	-0.483***	-0.449***	-0.522***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Top third baseline test scores	0.488***	0.431***	0.661***	0.535***	0.694***	0.570***	0.559***	0.525***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
TFA*Missing baseline test scores	-0.055	0.026	0.000	0.000	0.000	0.000	0.000	0.000
	(0.028)	(0.029)	(.)	(.)	(.)	(.)	(.)	(.)
Black	-0.331***	-0.318***	-0.195***	-0.227***	-0.198***	-0.230***	-0.295***	-0.262***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Hispanic	-0.081***	-0.123***	-0.049***	-0.073***	-0.045***	-0.089***	-0.076***	-0.057***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Asian	0.139***	-0.027***	0.207***	0.018***	0.254***	0.005	0.139***	0.091***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.006)	(0.005)	(0.005)
American Indian	-0.161***	-0.187***	-0.106***	-0.121***	-0.126***	-0.152***	-0.168***	-0.184***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.008)	(0.008)	(0.007)
Other Race	-0.129***	-0.113***	-0.069***	-0.054***	-0.068***	-0.047***	-0.058***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
Female	-0.079***	0.041***	-0.017***	0.036***	-0.008***	0.129***	-0.109***	-0.219***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.069***	0.083***	0.063***	0.090***	0.020***	0.080***	0.050***	0.050***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed: College or more	0.136***	0.146***	0.160***	0.180***	0.071***	0.172***	0.104***	0.122***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed: Missing	-0.039***	0.004	0.017***	0.026***	0.103***	0.078***	0.170***	0.202***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Student has disability	-0.336***	-0.419***	-0.248***	-0.318***	-0.169***	-0.390***	-0.191***	-0.153***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Disability status is missing	0.000	0.000	0.000	0.000	-0.102**	-0.458***	-0.075*	-0.248***
	(.)	(.)	(.)	(.)	(0.032)	(0.036)	(0.036)	(0.031)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Middle tercile baseline test score (math or reading); Standard errors in parentheses; baseline scores for students in grades 3-8 are baseline test scores. For high school, baseline scores are 8th grade math (for HS math and HS science) or 8th grade reading (for HS social studies and HS English).

Table 3.7 continued

Impact of TFA on same-year tests based on baseline proficiency, with school-subject-year fixed effects

	Elementary Math	**Elementary Reading**	**Middle Sch. Math**	**Middle Sch. Reading**	**High School Math**	**High School English**	**High School Science**	**High Sch. Soc. Studies**
Student is gifted	0.792***	0.709***	0.512***	0.452***	0.383***	0.422***	0.402***	0.407***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Gifted status missing	-0.681***	-0.457***	-0.623***	-0.664***	-0.320***	-0.063	-0.148	-0.344***
	(0.032)	(0.050)	(0.033)	(0.055)	(0.063)	(0.098)	(0.108)	(0.076)
Student ever IDed as LEP	-0.158***	-0.306***	-0.065***	-0.191***	-0.034***	-0.250***	-0.170***	-0.158***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)
LEP status missing	-0.257***	-0.325***	-0.441***	-0.482***	0.056	-0.776***	-0.283***	-0.267***
	(0.026)	(0.028)	(0.027)	(0.031)	(0.053)	(0.076)	(0.079)	(0.058)
Student over age for grade	-0.229***	-0.220***	-0.198***	-0.201***	-0.236***	-0.284***	-0.283***	-0.280***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Student under age for grade	0.048***	0.034***	0.065***	0.032***	0.129***	0.095***	0.175***	0.170***
	(0.006)	(0.006)	(0.004)	(0.005)	(0.005)	(0.007)	(0.006)	(0.006)
Student has repeated grade	-0.952***	-0.978***	-0.536***	-0.561***	-0.570***	-0.478***	-0.762***	-0.702***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.007)	(0.006)
Repeater status missing	-0.750***	-0.934***	-0.664***	-0.976***	-0.650***	-0.726***	-0.793***	-0.678***
	(0.191)	(0.192)	(0.090)	(0.098)	(0.036)	(0.050)	(0.042)	(0.036)
Constant	0.147***	0.119***	-0.041***	0.041***	-0.168***	0.040***	0.050***	0.111***
	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
R-squared	0.491	0.459	0.586	0.534	0.398	0.558	0.388	0.424
Observations	1963556	1950920	1911744	1905340	1385801	676912	1039382	1182520

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Middle tercile baseline test score (math or reading); Standard errors in parentheses; baseline scores for students in grades 3-8 are baseline test scores. For high school, baseline scores are 8th grade math (for HS math and HS science) or 8th grade reading (for HS social studies and HS English).

Table 3.8

Impact of TFA on same-year tests based on baseline proficiency and school level, with school-subject-year fixed effects

	Math	**ELA**
TFA	**0.052****	-0.030
	(0.020)	(0.020)
TFA*Middle School	-0.006	**0.052***
	(0.023)	(0.023)
TFA*High School	0.043	**0.058***
	(0.023)	(0.026)
Bottom Tercile baseline test (math or reading)	-0.582***	-0.597***
	(0.001)	(0.001)
Top Tercile baseline test (math or reading)	0.595***	0.498***
	(0.001)	(0.002)
TFA*Bottom tercile baseline test	0.012	0.020
	(0.022)	(0.023)
TFA*Top tercile baseline test	0.040	**0.113*****
	(0.030)	(0.031)
TFA*missing baseline test	**-0.062***	0.022
	(0.027)	(0.028)
Middle School*Bottom tercile baseline test	0.071***	0.025***
	(0.002)	(0.002)
Middle School*Top tercile baseline test	0.039***	-0.011***
	(0.002)	(0.002)
High School*Bottom tercile baseline test	0.184***	0.124***
	(0.002)	(0.002)
High School*Top tercile baseline test	-0.002	**0.018*****
	(0.002)	(0.003)
TFA*Middle School*Bottom Tercile baseline test	-0.003	-0.023
	(0.026)	(0.026)
TFA*Middle School*Top Tercile baseline test	0.018	-0.016
	(0.035)	(0.035)
TFA*High School*Bottom Tercile baseline test	0.028	0.026
	(0.027)	(0.031)
TFA*High School*Top Tercile baseline test	-0.038	**-0.115****
	(0.038)	(0.042)
Black	-0.248***	-0.268***
	(0.001)	(0.001)
Hispanic	-0.062***	-0.097***
	(0.002)	(0.002)
Asian	0.205***	0.001
	(0.002)	(0.002)

* p<0.05 ** p<0.01 *** p<0.001; Baseline tests are lagged test scores in grades 4-8 and 8th grade tests in grades 9-12; Omitted categories are: White; Parent Ed. High Sch. or less; Less than median baseline test score, Elementary School; Standard errors in parentheses.

Table 3.8 continued

Impact of TFA on same-year tests based on baseline proficiency and school level, with school-subject-year fixed effects

	Math	**ELA**
American Indian	-0.134***	-0.155***
	(0.003)	(0.003)
Other Race	-0.092***	-0.080***
	(0.002)	(0.002)
Female	-0.038***	0.052***
	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.048***	0.081***
	(0.001)	(0.001)
Parent Ed: College or more	0.106***	0.153***
	(0.001)	(0.001)
Parent Ed: Missing	0.048***	0.037***
	(0.001)	(0.002)
Student has disability	-0.270***	-0.376***
	(0.001)	(0.001)
Disability status is missing	-0.034	-0.431***
	(0.031)	(0.039)
Student is gifted	0.584***	0.560***
	(0.001)	(0.001)
Gifted status missing	-0.596***	-0.491***
	(0.022)	(0.034)
Student ever IDed as LEP	-0.098***	-0.256***
	(0.002)	(0.002)
LEP status missing	-0.301***	-0.421***
	(0.018)	(0.020)
Student over age for grade	-0.219***	-0.221***
	(0.001)	(0.001)
Student under age for grade	0.091***	0.048***
	(0.003)	(0.003)
Student has repeated grade	-0.687***	-0.639***
	(0.002)	(0.002)
Repeater status missing	-0.653***	-0.813***
	(0.032)	(0.045)
Constant	-0.026***	0.069***
	(0.001)	(0.001)
R-squared	0.497	0.501
Observations	5261101	4533172

* p<0.05 ** p<0.01 *** p<0.001; Baseline tests are lagged test scores in grades 4-8 and 8th grade tests in grades 9-12; Omitted categories are: White; Parent Ed. High Sch. or less; Less than median baseline test score, Elementary School; Standard errors in parentheses.

Chapter 4

Teacher Quality and Teach For America's Increasing Effectiveness over Time

Teach For America (TFA) is a selective alternative teacher certification program that has a growing influence on the way teachers are recruited and trained in schools across the country. Although 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),[27] 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 evaluated.

A growing body of work has evaluated TFA's effects at different points in time. Early evidence on TFA's effects was mixed, but a number of recent studies find positive effects of TFA. 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, growing program 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 effectiveness in 2005 (Jaramillo, 2002; Sawchuk, 2009; Tourangeau, 2003). Subsequently, TFA initiated a second period of expansion which it plans to complete in 2015 (Kopp, 2011). However, no study to date has tested whether TFA teacher effectiveness has changed over time as TFA modified its program in an effort to improve student outcomes. Evidence from Chapter 3 and other work examining TFA across subject areas suggests that TFA is not uniformly effective in all subjects and grade levels. This chapter takes a similar approach

to examine whether TFA's impact on student achievement has changed over time, paying specific attention to how changes may vary across subjects and grade levels.

One of the primary qualities that differentiates TFA from traditional certification pathways is that it seeks to recruit individuals from more selective backgrounds into teaching. This suggests that differences in observable teacher quality between TFA and non-TFA teachers might explain differences in the effects of TFA teachers. Evidence from North Carolina confirms that high-poverty schools, including those where TFA teachers are placed, have more inexperienced teachers from less competitive undergraduate institutions, with below-mean licensure test scores and fewer teachers with National Board Certification (Clotfelter et al., 2007b, 2006).

Differences in teacher quality may also explain any changes in TFA teachers' relative effectiveness over time. Changes in the effects of TFA over time may result from TFA teacher quality improving as the program becomes more selective in its own recruitment. Alternatively, TFA teachers may look comparably better or worse without changing at all if counterfactual teacher quality declines or improves during the same time period. Counterfactual teacher quality might be improving in similar ways to TFA teacher quality because North Carolina's own alternative teacher certification programs might be drawing more selective individuals into teaching. Alternatively, North Carolina's traditional teacher certification programs may also be improving their selectivity and quality, which could also lead TFA teachers to look less or similarly effective over time. In contrast, TFA teachers may look better over time if counterfactual teacher quality in TFA schools declines over the same time period. Thus it is possible that the transfer of higher-quality teachers out of such schools in greater numbers over time might contribute to those schools seeking out TFA teachers.

This study examines the effect of TFA across elementary, middle, and high school using a unique dataset from the state of North Carolina, spanning the years 2000 - 2011. These data allow me to test whether TFA's impacts have changed across the pre- and post-2005 period, as North Carolina experienced an expansion that mirrored TFA's national expansion. In addition, I examine whether teacher quality explains (1) any main effects of TFA or (2) changes in TFA's effects over time.

I find evidence that TFA's effectiveness has increased over time in several subjects and grade levels relative to non-TFA teachers, specifically in middle school math, high school science, and high school social studies. This pattern of improvement is shared more broadly across several other subjects and grades, although the differences are only marginally significant. 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. One notable exception is high school math, where, if anything TFA teachers appear to be losing ground relative to their non-TFA counterparts.

I also find that the main effects of TFA on student achievement are not explained by most observable teacher quality indicators. The only exception is teacher Praxis scores, which account for approximately one third of TFA's effect on student achievement across subject areas and grades. Likewise, none of the other measures of teacher quality explain TFA's improvements after 2005, except for average Praxis scores, where teachers' Praxis scores account for approximately a quarter of the improvement in TFA effects over time in the three subjects in which TFA's effects grew.

Together these results suggest that TFA's effects are shared across most subjects and grades, and that TFA has improved over time in some subjects, but not others. Both TFA's

effects on student achievement and changes in effectiveness as the program has matured are partially explained by teacher Praxis scores, but not other measures of observable teacher quality. This suggests that TFA's effects on student achievement and improvements over time are also due to other aspects of the TFA training program or unobservable characteristics of TFA teachers that cannot be captured by traditional measures of teacher quality. Finally, they underscore TFA's widely identified difficulties in improving student reading achievement in the elementary grades.

Literature Review

Teach For America – Background and Effectiveness

It is important to evaluate TFA's effectiveness, not only to learn whether it improves student performance in underserved schools, but also because TFA makes a provocative argument about how teacher selection and training are related to teacher effectiveness and quality. From its inception, TFA has 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 be effective with minimal training or teaching experience. Even more controversial, TFA promotes the idea that teacher quality can be measured, and that high quality teachers can be identified through a rigorous selection process that seeks out personal qualities without evidence of specific teaching skill. If TFA is at all effective in improving student achievement, this suggests that alternative pathways to teacher certification are a promising complement to existing teacher preparation programs. In contrast, if TFA has a null or negative impact, its use and the proliferation of such practices to identify teachers should be reconsidered.

The best experimental and quasi-experimental evidence suggests that TFA teachers are more effective on average than other teachers in math and science, but not in reading at the elementary and high school levels (Clark et al., 2013; Glazerman et al., 2006; Kane et al., 2008; Xu et al., 2011). Quasi-experimental and descriptive studies using administrative data yield mixed conclusions, with some finding positive effects of TFA teachers relative to non-TFA teachers, others finding a varied pattern by subject, and still others finding that TFA teachers performed no differently or worse than non-TFA teachers (Darling-Hammond et al., 2005; Schoeneberger et al., 2009; Ware et al., 2011). While these studies represent TFA impacts ranging across the years 1997 to 2011, the longest time span covered an individual study is only six years and does not include high school, and as a result none of these studies test whether TFA's effects are consistent over time, particularly in the subject areas where TFA's effects have been shown to be largest in other work (Darling-Hammond et al., 2005; Dobbie, 2011).

Although the extant research examining TFA's effects has not examined whether TFA's effects have changed over time, TFA as an organization has undergone many substantial transitions since its inaugural year in 1990. TFA underwent a period of expansion in the early 2000s, which has been followed with several additional waves of expansion. It has grown substantially in size and national presence, expanding from five original sites, including one in Eastern North Carolina, to approximately 11,000 corps members in 48 regional sites in the 2013-14 school year (teachforamerica.org, 2013c). In the late 2000s TFA continued to pursue growth within existing regions and into new regions, with a goal of placing 15,000 corps members in 60 regions by 2015, all while building its alumni network and educational leadership capacity (Kopp, 2011; Sawchuk, 2009; teachforamerica.org, 2014).

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 worked 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 to measure teacher effectiveness internally, which was also incorporated into its teacher training model in 2005. Through this training reform, TFA began to define and measure its success in terms of improvements in student achievement. 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, it has worked to address critiques of TFA's culture by increasing the racial and socio-economic diversity of its corps and by recruiting more individuals with graduate degrees and post-baccalaureate professional experience (Jennings, 2014; teachforamerica.org, 2013c).

Indicators of Teacher Quality

TFA critics often argue that TFA teachers must be less effective because they lack certain types of formal qualifications and are inexperienced. While the literature on teacher quality suggests that teacher experience has a positive effect on student achievement, at least when

moving from no experience to some (Clotfelter et al., 2007b, 2010), many other observable measures of teacher quality, such as having a masters degree, are inconsistently related to teacher effectiveness (Clotfelter et al., 2010). In addition, while TFA detractors often champion the quality of non-TFA teachers, it is not clear that their quality is particularly high. Research suggests that 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 quality, 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 et al., 2004). This suggests that new teachers from higher-caliber backgrounds are becoming increasingly rare.

Further, studies that find negative effects of TFA often compare TFA to teachers across the district or region, who are often teaching different populations of students, rather than looking specifically at teachers in the same under-resourced schools where TFA places teachers (Darling-Hammond et al., 2005; Laczko-Kerr & Berliner, 2002). As TFA seeks to staff hard to fill positions in underserved districts, these schools may have particularly weak teachers. This is because staffing is typically done by seniority, and as teachers gain experience, they become increasingly likely to leave low-performing, high-minority schools in lower-income areas (Carroll et al., 2000; Cha & Cohen-Vogel, 2011; Cohen-Vogel et al., 2013; Cohen-Vogel & Osborne-Lampkin, 2007; Darling-Hammond, 2003; Esch et al., 2005; Hanushek et al., 2004a; Scafidi et al., 2006), so that non-TFA teachers in TFA placement schools are likely less qualified and less experienced than non-TFA teachers in more affluent, higher-achieving schools. Evidence from North Carolina confirms that high-poverty schools, including those where TFA

teachers are placed, have more inexperienced teachers from less competitive undergraduate institutions, with below-mean licensure test scores and fewer teachers with National Board Certification than elsewhere in the state (Clotfelter et al., 2007b, 2006).

However, while non-TFA teachers in TFA placement schools come from weaker backgrounds than TFA teachers and non-TFA teachers in more affluent schools on average, the quality of these counterfactual teachers may vary considerably, which could be the result of several different factors. One potential source of variation in teacher quality across school level and subject is differences in the academic background of non-TFA teachers, which might contribute to their skills as a teacher. It may be the case that non-TFA teachers at particular stages have more rigorous educational backgrounds and experience that make them stronger teachers than the non-TFA teachers in other school levels. For example, teachers with higher level math skills typically teach high school math (Boyd et al., 2012), and most high school math teachers have several years of college-level math coursework and many majored in math or a quantitative subject (Ingersoll, 1999). However, as has been documented in previous research, few elementary school teachers take advanced math or courses in college (Book & Freeman, 1986), and few of them major in math, science or similar subjects (Ingersoll & May, 2012). Thus, in high school math, the non-TFA teachers' instruction might compare favorably relative to that of TFA teachers, while in elementary school, TFA teachers' math instruction might be substantially better than that of non-TFA teachers. While the TFA teachers themselves might have similar levels of math experience regardless of school level, the counterfactual teachers may differ dramatically in their relative level of experience and comfort with math across school levels. This may lead TFA teachers to be more effective in elementary math than high school math, simply by virtue of the weaker relative math skills of the counterfactual teacher.

It may instead be the case that traditional teacher training programs are more effective at training teachers for certain subjects and grade levels. That is, if traditional programs have a strong conception of what an effective high school math teacher needs to know, but little sense of what an elementary language arts teacher needs to master, then we would expect high school math training to produce more successful teachers than elementary language arts training. Research in North Carolina finds some evidence of this. For example, Clotfelter, Ladd, & Vigdor (2010) find that receiving a master's degree has a small positive relationship with student achievement in high school, while it has a negative relationship in elementary school. Likewise, if state certification boards have a better understanding of what teachers need to know in order to be successful teachers in some subjects than others, then state certification exams in some subjects might also be more predictive of student outcomes. Evidence suggests that training, certification, and student learning are more closely aligned in some subjects than others. In North Carolina, high school teachers' certification test scores in math and science are positively associated with their students' scores on end-of-course tests in the same subjects, with math exhibiting a larger effect. Interestingly, their results suggest that Praxis scores matter more in high school than in elementary, which might also suggest that content knowledge matters more at higher grade levels (Clotfelter et al., 2007b, 2010).

Although stronger teacher backgrounds have a positive impact on student achievement, they are also associated with greater levels of teacher turnover, which may also contribute to TFA impacts. Previous research suggests that teacher turnover is greater among teachers with higher test scores and that high school chemistry and physics teachers are more likely to leave than teachers in other content areas (Murnane & Olsen, 1989, 1990). A meta-analysis of teacher turnover confirms these findings. Borman and Dowling (2008) find that high school math and

science teachers have higher rates of turnover than elementary school teachers, who in turn have higher rates of turnover than high school teachers of non-math and non-science subjects. Little empirical evidence exists about the level of turnover and difficulty of staffing middle schools, however anecdotal evidence suggests that middle school grades are the most difficult to staff and are plagued with teacher turnover (Useem & Neild, 2001), and a survey of education majors from North Carolina and Virginia suggests that middle schools are the least desirable due to concerns around discipline and adolescents' attitude problems (M. S. Carter & Carter, 2000). This suggests that TFA teachers may face relatively inexperienced peers in high school STEM subjects and middle school than in other grade levels and subject areas, which may also impact their relative effectiveness.

Hypotheses About Where Experience Might Matter Most and Why

The previous literature on the impacts of observable measures of teacher quality suggests that factors such as years of experience and Praxis certification scores are likely to have a larger impact on student achievement than other measures of teacher quality, both of TFA and non-TFA teachers. Likewise, these observable measures of teacher quality seem the most likely to be able to explain any of the effects of TFA or any changes in its effectiveness over time. It also seems likely that controlling for measures of teacher quality 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 gain experience and additional credentials.

In terms of predicting which subjects will be most impacted by teacher quality, the literature is less clear. Following the findings that TFA effects were largest in high school math and science, and larger in math than ELA more generally, it seems likely that observable

measures of teacher quality play the largest role in explaining TFA effectiveness and change over time in STEM subjects. Given that TFA typically places teachers in hard-to-staff schools, where veteran teachers are more likely to transfer to other schools and teachers of all experience levels are more likely to quit, these schools are also more likely to have inexperienced teachers in general, but particularly in hard to staff subjects. This suggests that non-TFA teachers in math and science in high school that employ TFA teachers, as well as middle schools with TFA teachers, are more likely to be novices than elementary school teachers and high school English teachers. If experience is an important factor, this could lead to a larger TFA advantage in high school math and science and middle school. However, these counterfactual teachers are also most likely to have higher test scores than other non-TFA teachers, making them potentially more effective and offsetting their relative degree of inexperience. Thus, while it seems likely that teacher quality measures have the largest impacts on high school math and science, other evidence conflicts with this prediction.

Data

I examine the relationship between TFA, student achievement, and observable teacher quality over time using administrative data from the state of North Carolina. TFA places teachers in two 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 (teachforamerica.org, 2013b).

Administrative data from North Carolina were provided by the North Carolina Education Research Data Center (NCERDC) at Duke University. NCERDC provided demographic and assessment records for students which were matched to their classroom teachers.[28] These data are well suited to the examination of the effectiveness of TFA versus non-TFA teachers. In addition

to the prior studies that directly compare TFA teachers to non-TFA teachers (Henry et al., 2010; Xu et al., 2011), these data have also been used for numerous studies examining teacher quality (Clotfelter et al., 2006), the feasibility and stability of teacher value-added models (Clotfelter et al., 2007b, 2010; Goldhaber & Hansen, 2010; Rothstein, 2009, 2010), and teacher persistence and turnover (Goldhaber et al., 2011).

All North Carolina TFA teachers (also called corps members) from the entering corps of 1999 to 2010[29] 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).[30] 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.[31] 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 corps members placed in Eastern North Carolina each year more than doubled. TFA corps members by placement region and year for grades K-12 in 2006-2011 are shown in Table 4.1.

[Insert Table 4.1 Here]

From 2000-2011 TFA teachers taught in 435 schools in 74 LEAs and charter schools in North Carolina, primarily concentrated in the north eastern corner of the state or near the Charlotte metro area. In this study, 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 directly the effects of having a TFA teacher relative to other teachers the same students could have had for the same courses. I describe this approach in more detail in the methods section below, but to provide more detail about how students and teachers in these two groups differ, a comparison of the demographic characteristics of the students with a TFA teacher versus four different comparison groups is shown in Table 4.2. 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 case, because these students had the potential to be assigned to a TFA teacher in the same year, but were not, and thus it is used to identify counterfactual students and teachers in the analyses described below. The analytic sample throughout this study consists of only those students with TFA teachers in group 1, where there was the potential of being assigned to a TFA teacher in group 2.

[Insert Table 4.2 here]

As indicated in Table 4.2, students of TFA teachers are most similar to other students in non-TFA classrooms in the same schools. They have fewer parents with college degrees or above, more of them are African American, fewer of them were ever identified, and more of them are over-age for grade or have repeated a grade than students in other schools and LEAs.

They also have more students from the bottom third of the baseline test score distribution than teachers in the other groups.

There are also differences in the observable teacher quality 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), only about half are fully certified compared with three-fourths of non-TFA teachers, and fewer than five percent of them have a master's degree or more, compared with at least 25 percent in the other groups. However, TFA teachers' mean Praxis scores are .35 standard deviations higher than those of non-TFA teachers, and more than .4 standard deviations higher than those of the counterfactual teachers in the same schools, grades (subjects), and years in which TFA teachers are teaching.[32] This suggests that, while TFA teachers lack the experience and formal credentials of the non-TFA teachers, they may have a stronger grasp on the material the certification board deemed important for teaching, despite limited formal training.

The outcome of interest in this study is end-of-grade or end-of-course student test scores. In grades three through eight, tested subjects include math and reading, which I combine across grades 3-5 and 6-8 for parsimony.[33] In high school, tested subjects include math, English, science, and social studies. Three of these four global subjects consist of several individual tested subjects, all of which are combined for 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 consist of Economy, Legal, and Political Systems, US History, and Civics and Economics; and English consists of English 1. 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 in which there was no social studies test given.

Student test scores are standardized within year and grade for end-of-grade tests and within year and subject for high school end-of-course tests. 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. Lagged test score controls are included in models examining outcomes in grades three through eight. 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 the student controls listed in Table 4.2, and several of the models also examine the role of observable teacher quality in predicting student achievement using the teacher quality measures shown in Table 4.2. In an effort to minimize missing data, the student characteristics were filled in from across several individual datasets provided by the NCERDC. For student gender and race ethnicity, the modal value for an individual across these datasets was taken and filled in cases that it was missing. For example, if student gender was present in some years and missing others, or if it agreed in three out of four total years the student was in the dataset, the modal non-missing value was taken for all four years.

Using the data filled in across several data sources, I created the 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 taken. Parental education 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 they were never identified as having any of these three characteristics. 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. The student the over-and under-age variables are not fixed from year to year and can change following a retention or early promotion, unlike the rest of the student characteristics which are filled in across multiple years of student observations.

North Carolina began compiling a database of teacher quality measures in 1995. Thus although the student data begin in 2000, I include teacher data beginning in 1995. Teacher characteristics are filled in using a similar procedure to that of the student characteristics. The majority of the observable teacher quality 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. This threshold is identified to be consistent with the cutoff used in Study 1 which draws its cutoff from Glazerman et al. (2006). Teacher certification is included as an indicator where a one indicates that a teacher is fully certified. Teacher education is examined using an indicator for whether a teacher has a master's degree or higher (yes = 1; no = 0). I created a variable for undergraduate selectivity by matching students' undergraduate degree

granting institutions to the 2009 Barron's Selectivity rankings, which include four categories (1 = most competitive; 2 = highly competitive plus; 3 = highly competitive; 4 = 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 quality. The final teacher quality measure is average Praxis licensure test score. Though the vast majority of teachers had a single Praxis score, this variable takes the average of all Praxis scores a given teacher has taken. Each individual Praxis score is standardized within test type and the year in which the test was taken. These standardized values are then averaged within the individual teacher.[34]

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 quality 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 use a version of teacher value-added models with fixed effects for school, grade, and year combinations.[35] 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 primary evaluation of achievement gains in TFA versus non-TFA classrooms will be 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 (c.f. Table 4.2).

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:

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

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 linear function of the student's prior-year test score, 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. 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 grouping makes comparisons within school-grade-year units so that students who have TFA teachers are compared only to 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 eliminates 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. It also eliminates 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 grade level that may be due to unobservables related to the student, the teacher, or parental preferences. The presence and magnitude of this bias in North Carolina and other settings continues to be debated (Clotfelter et al., 2006; Goldhaber & Hansen, 2010; Koedel & Betts, 2011; Rothstein, 2009), however this approach continues to be used in many administrative data applications for comparing the impact of TFA teachers on student achievement (e.g., Boyd et al., 2009, 2006). Initial models examine the relationship between having a TFA teacher and student test scores as described above. Subsequent models add controls for teacher quality to examine whether the main effects of TFA can be explained by observable teacher quality. I then examine whether the effect of TFA changes over time by including an interaction with TFA and the post-2005 period. While the school-grade-year or school-subject-year fixed effects cause the main effect of the post-2005 period to drop out of the model due to co-linearity, the TFA*post-2005 term remains in the model and this term reports whether the effect of TFA differed pre- and post-2005. Finally, to test whether or not teacher quality can explain any changes in the relationship between TFA and student achievement, I estimate models with terms for teacher quality, as well as the interaction of each of the teacher quality measures and the post-2005 period.

Results

Before testing whether the relationship between TFA and student achievement (1) varies over time and (2) is explained by observable teacher characteristics, I first examine the main effects of TFA teachers on student achievement across school levels and subjects. This initial analysis replicates what is presented in Study 2. These results are shown in Table 4.3.

[Insert Tables 4.3 here]

Table 4.3 shows that the effects of TFA are positive in most subjects and across most grade levels, although they differ in magnitude. TFA effects are by far the largest in high school science, where having a TFA teacher improves student achievement by 0.197 standard deviations. The TFA effect in math is also fairly substantial at 0.112 standard deviations. Effects in high school English, social studies, and elementary and middle school math are all similarly sized, ranging from 0.036 standard deviations to 0.050 standard deviations. The one exception to these positive effects is elementary reading, where the TFA coefficient is negative and marginally significant.

Additional models testing whether these broad patterns hold for individual subjects and grades are shown in Appendix Tables 4.1 and 4.2. These results show that across all elementary and middle school grades, TFA has a positive effect in math. In contrast, the results across grades are inconsistent for elementary and middle school reading, where the TFA effect is positive in 8th grade, negative in 5th grade, and indistinguishable from zero in the remaining three grades. Looking across subjects in high school, TFA has positive effects in every high school math subject (with the largest effects (0.163 SD) in geometry), and in three of the four science subjects (with the largest effects (0.355 SD) in chemistry). Results are more mixed in social studies, with large negative effects in Economics, Legal, and Political Systems[36] (-0.971 SD), and moderate positive effects in the other two subjects. There is only one English test in high

school, rendering further examination unnecessary. Thus, overall, the results from the individual subject and grade level models generally conform to the broader results presented in Table 4.3, with some notable exceptions.

I next examine whether the relationship between TFA teachers and student achievement, which was widely shared across most subjects and grades, can be 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, and the selectivity of the teachers' undergraduate institution are shown in Table 4.4. These results present only the coefficients for having a TFA teacher and the teacher quality measure of interest. Full results from these models are reported in Appendix Tables 4.3 – 4.8.

[Insert Table 4.4 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 quality 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 4.4., the coefficients for TFA remain significant and increase in every subject and grade level, sometimes by more than half. Comparing models 4.3 and 4.4 thus indicates that TFA teachers perform better than their counterfactual teachers even though they have lower levels of typical measures of teacher quality, like experience. Thus, when we control for the differing levels of experience between

TFA and non-TFA teachers, in essence comparing TFA teachers to counterfactual teachers with similar levels of experience, we find that TFA teachers do even better than non-TFA teachers.

The one observable measure of teacher quality that does explain part of the relationship between TFA and student achievement is teacher Praxis scores. Including controls for average Praxis scores reduces the magnitude of most of the TFA coefficients by 10-25 percent. In addition, adding the Praxis control also increases the absolute magnitude of the significant, negative coefficient on Elementary reading, so that TFA teachers have a negative effect relative to teachers with similar average levels of Praxis scores of -0.023 standard deviations. 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.

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, which is shown in Appendix Table 4.9. These models confirm the results shown in Table 4.3., 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. Overall, the coefficients for the veteran teacher comparisons are of similar magnitude to those comparing TFA teachers to all non-TFA teachers.

I next examine whether the effect of TFA changes as the TFA program has matured (Tables 4.5 and 4.6). These models replicate the regression results for the full 2001-2011 period shown in Table 4.3. Then, in two additional columns per outcome, they show results restricted to the years 2001-2005 and 2006-2011.

[Insert Tables 4.5 & 4.6 here]

Table 4.5 shows models examining TFA's effects in the less- and more-mature TFA periods in elementary and middle school. The coefficients for elementary math, and middle school math and reading follow a similar pattern. In these three subjects, the overall effect of TFA is positive and significant for the full panel, the coefficients for TFA are small and insignificant in the 2001-2005 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 suggestive of an improvement in the effect of TFA from the earlier to the later period. In elementary reading, a similar trend is present, however, instead of being positive overall, the effect of TFA on elementary reading is not different from zero for the full panel. While in 2001-2005 TFA had a negative effect of 0.086 standard deviations on elementary reading, by the later period TFA teachers had improved to catch up to a level that is no different from that of non-TFA teachers.

Table 4.6 shows models examining the same transition from earlier to later TFA for the four high school subjects. High school English follows the same pattern seen in three of the earlier subjects where it has a significant positive effect overall, in which TFA teachers were no different than non-TFA teachers in the earlier years and become more effective in the later TFA period. High school math and science follow a different pattern. In both of these subjects, TFA began positive in the earlier time period and got even better as the TFA program matured, so that by the later period TFA science teachers were more than two tenths of a standard deviation better than non-TFA science teachers. In high school social studies, TFA effects also appear improve substantially over time. Here TFA teachers were substantially worse than non-TFA teachers in the early period (0.35 standard deviations) and in the later period have improved to be better than non-TFA teachers by 0.063 standard deviations.

While the results in Tables 4.5 and 4.6 both suggest that TFA effects are improving over time in all subjects, these models do not provide a formal test of whether or not TFA effects are significantly different across the two time periods. Models including an interaction term with TFA*Post-2005 are shown in Table 4.7 to test whether the differences between the two time periods are significant.

[Insert Table 4.7 here]

The results in Table 4.7 show that some of the patterns shown in tables 4.5 and 4.6 are statistically significant while others are not. The subjects with a significant difference between the pre- and post-2005 period are middle school math and high school science and social studies. The magnitude of the difference is fairly small in middle school math (0.038 standard deviations). Notably, I find that the main effect of TFA, which represents the effect of TFA in the earlier period, is no longer significant, congruent with the results from previous models. In high school science, both the main effect of TFA and the interaction are significant suggesting that TFA had a positive effect in the earlier period, and improved by 0.091 standard deviations in the post-2005 period. Social studies also shows signs of significant and substantial improvement between earlier and later TFA. The effect of TFA in the earlier period is large and negative at 0.473 standard deviations. By the later period this negative effect of TFA is gone, as the TFA*Post-2005 coefficient is 0.539. These changes appear to be driven by particularly large negative effects on achievement in the Economics, Legal, and Political Systems courses in the early period and modest positive effects on achievement in other social science subjects in the later period. Given the changes in the tested subjects in social studies over time, these results should be judged with a greater degree of caution than results from the other subject areas and grade levels.

Notably, the coefficients for the earlier TFA period are significant for high school math, where it is positive 0.158 standard deviations, and elementary reading, where it is -0.087 standard deviations, but in both of these cases, the coefficient on the timing interaction is only marginally significant, suggesting that there may in fact be some change over time that is positive for elementary reading and, if anything, negative for high school math. Neither the interaction terms nor the main effects are significant in the models for elementary math, middle school reading, and high school English. Appendix Table 4.10 replicates this analysis, but compares TFA teachers to veteran non-TFA teachers only. In these models, the timing coefficients are significant and of a similar magnitude for the same three subjects as the comparisons made relative to all non-TFA teachers, suggesting the TFA improvement occurs regardless of the experience level of the counterfactual teachers.

Finally, in Table 4.8, I examine whether these changes over time can be explained by differences in teacher quality. This table summarizes the key coefficients from individual models for each measure of teacher quality.[37] The corresponding full models are shown in Appendix Tables 4.11 – 4.16.

[Insert Table 4.8 here]

These results mirror those showing the relationship between the main effect of TFA and teacher quality. The majority of the teacher quality measures have no impact on the interaction of TFA and the post-2005 indicator, with the exception of mean Praxis scores, which reduce the magnitude of these coefficients by 10 to 20 percent. Supplemental analyses (not shown), suggest that TFA and non-TFA Praxis scores have converged between 2000 and 2011. Analyses examining whether TFA Praxis scores or counterfactual teacher Praxis scores were more salient in explaining these differences between early and later TFA were inconclusive. However, some

evidence suggests that the correlation between Praxis scores and student achievement for TFA teachers decreases in the later time period, even as the results in Tables 4.5 and 4.6 show that the effects of TFA get stronger over time in some subjects. These findings further suggest that program components, rather than teacher quality as it is often defined by districts and schools, are driving TFA's improvements over time.

Conclusion

Chapters 2 and 3 in this dissertation 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. Building on these examinations of student-driven treatment heterogeneity, this chapter examines whether variation in the effects of TFA can be explained by teacher quality and the evolution of the TFA program.

TFA was founded with the assumption that recruiting highly selective individuals into the classroom would improve student outcomes. As the program has evolved and grown, it has become ever-more focused on improving its selection and training processes as a means to further improve student outcomes. The literature on the relationship between teacher qualifications and student outcomes is decidedly mixed (Clotfelter et al., 2010; Goldhaber & Anthony, 2007; Leigh, 2010; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004), but finds strong evidence that teachers of higher quality are unequally distributed across schools, and are particularly lacking in low-income schools (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2008; Boyd et al., 2005a; Boyd, Lankford, Loeb, & Wyckoff, 2005b). While the existing literature suggests that TFA teacher backgrounds are different than those of non-TFA teachers, it was not clear whether these differences in teacher quality explain TFA teachers' differential effectiveness in most subjects and grade levels. This study provides evidence suggesting that

most of the commonly used measures of observable teacher quality cannot explain differences between TFA and non-TFA teachers. 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 quality (e.g full certification, advanced degrees, experience, etc.). Thus, accounting for these differences only increases the relative effectiveness of TFA teachers. This is noteworthy, and given past criticisms of TFA centered on experience, it is likewise important to note that when I compare TFA teachers only with experienced teachers, I find that TFA teachers outperform even experienced non-TFA teachers.

The only observable teacher quality measure that seems to explain part of the TFA effect is average teacher Praxis score. Praxis scores not only explain away part of the relationship between TFA teachers and student achievement, but they also partially explain the larger effect of TFA in the post-2005 period. However, the controls for Praxis scores do not fully account for the majority of TFA main effects or the changes observed over time. This suggests that something besides changes in observable teacher quality using these measures is driving TFA's effects.

An alternative explanation for why TFA teachers are outperforming non-TFA teachers draws from the evolving nature of the TFA program itself. Rather than changing the quality of TFA recruits by finding individuals with higher standardized test scores or master's degrees, these results suggest that TFA might have made improvements to its training that resulted in TFA teachers becoming more effective as the program has matured. Of course it is also possible

that something about the TFA selection process is identifying difficult-to-observe characteristics that are leading to changes in TFA effectiveness.

Insofar as my results suggest that important aspects of the TFA selection and training process are not identifiable in administrative data, they suggest that we may need to reconsider how we select teachers and conceptualize teacher quality. While it is difficult to identify precisely the features of the TFA program that have changed over time without historical data from the organization itself, it is possible to examine whether TFA effects have changed over time. In addition, it is possible to test whether traditional measures of teacher quality can account for any observed changes over time. While it is likely impossible to definitively determine whether TFA training or more rigorous selection are responsible for changes in TFA's effectiveness, we can rule out some standard observable teacher characteristics as explanations.

As shown in Chapter 4, TFA teachers clearly outperform non-TFA teachers in most subjects. Likewise, TFA has improved over time in several of these subjects, but not others. In particular, high school science shines as an area where TFA teachers are more effective than non-TFA teachers. This appears to be especially the case in chemistry, biology, and physical science, and less so in physics. One average, TFA had a positive effect on science in the earlier time period and this effect increased as the program matured, suggesting science as one of the areas that benefits most from TFA teachers. TFA teachers also showed improvements in middle school math and high school social studies, although they were clearly much less effective at social science in the earlier period of TFA. This result is driven by some very large negative impacts in a few of the earlier years of TFA in high school social studies, making it difficult to interpret.

While overall this study suggests that TFA's effects are positive and have likely improved, results in two subject areas deviate from this trend. 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 the distributional results from the experimental TFA data in Chapter 2. Future work might examine whether TFA's effects on the distribution of student achievement in elementary reading in North Carolina mirror the results found using the Mathematica data, and whether improvements are shared across the distribution as TFA has matured over time.

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 significant at the .05 level. This anomalous change over time is curious, though it might indicate that non-TFA teachers in this area are beginning to incorporate many of the same practices that TFA teachers use. Subject-specific analyses (not shown) suggest that it is 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 is the first to examine descriptive changes in the relationship between TFA and student achievement over time. It examines whether TFA's effects have improved as the program matured throughout the 2000s, TFA's second decade of existence in North Carolina and nationally. While my results do not provide evidence of positive TFA effects across the board, they suggest that TFA teachers are better than the currently available alternative teachers who are willing to teach in the same schools and subjects/grades with only one clear exception, elementary reading. My results also suggest that, in several subjects, TFA's effects have grown over time. What my results cannot clearly explain is why. Measures of observable teacher quality do little to explain TFA main effects or changes over time, with the exception of Praxis scores, which explain about one third of the effect. Thus, in future work, I hope to work to identify the sources of these differences and change, potentially through observations of TFA and non-TFA classrooms, examining TFA training materials, or other insights gained by further conversations with the organization. Overall, my results suggest that the TFA program is more effective than its critics suggest, and that it is especially so in subjects which face high degrees of teacher turnover. This suggests that principals in similar schools would do well to seek out TFA teachers, particularly in the area of high school science, but may benefit less if they are concerned with struggling readers in elementary school.

[27] 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) .

[28] Technically, these 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. Clotfelter, Ladd, and Vigdor (Clotfelter et al., 2007a) examine the accuracy of these matches for elementary school in detail, also examining whether the teacher matched a personnel record and taught a valid math or reading course code in the relevant subject in the same year.

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. For parsimony, I currently refer to the test administrator as the teacher.

[29] 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.

[30] 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 2012/13 there were 115 LEAs and 107 charter schools (http://www.ncpublicschools.org/docs/fbs/resources/leacharterlist.pdf, 2013).

[31] 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 4.2.

[32] 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 sampled includes teachers that instruct untested subjects, which have lower average Praxis scores than those teaching tested subjects.

[33] 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.

[34] While ideally I would match using the subject the teacher was teaching, many teachers did not have scores in the subject their students were being tested in. The approach taken here thus allows those Praxis scores to be used.

[35] 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).

[36] It is important to note that the exams in Social Studies have 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. The remaining exam subjects have been mostly constant over time, but were updated in either 2005/2006 or 2006/2007 to increase their difficulty (North Carolina Department of Public Schools, 2011).

[37] Although it would be useful to examine the impacts of these measures of teacher quality controlling for the other measures in a model which included all of the measures together, when attempted, several of the coefficients for these measures dropped out of the model when entered simultaneously due to issues of co-linearity across several of the measures.

Table 4.1.

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

Year	Charlotte	Eastern North Carolina	Total
1999	0	28	28
2000	0	25	25
2001	0	26	26
2002	0	41	41
2003	0	7	7
2004	0	1	1
2005	54	43	97
2006	58	60	118
2007	55	51	106
2008	98	82	180
2009	63	56	119
2010	104	52	156
Total	432	472	904

Table 4.2

Student and teacher characteristics across comparison samples

	TFA Students	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 Gender						
Male	50.0	50.1	50.4	50.5	50.5	50.2
Female	50.0	49.9	49.6	49.5	49.5	49.2
Missing Gender	0.0	0.0	0.0	0.0	0.0	0.6
Parent Education Level						
High school or less	25.0	20.5	16.9	17.5	23.0	18.8
Some college/trade school	24.6	21.2	20.3	21.1	25.9	22.0
College degree or higher	23.7	35.3	48.0	42.0	34.8	40.8
Missing Parent Ed.	26.6	23.0	14.8	19.4	16.3	18.4
Ever identified as LEP						
Never identified	87.4	87.1	90.7	92.9	92.2	91.5
Ever identified	10.8	11.0	8.4	6.4	7.2	7.1
Missing LEP information	1.7	1.9	0.9	0.7	0.6	1.3
Student Ethnicity						
White	15.6	29.6	46.9	54.9	70.7	57.2
Black	67.5	52.1	39.1	30.4	17.5	28.5
Hispanic	10.6	11.1	7.8	7.1	7.0	7.3
Asian	2.3	3.1	3.1	2.5	1.5	2.3
American Indian	0.9	0.6	0.6	1.6	1.0	1.4
Other ethnicity	3.0	3.5	2.6	2.9	2.3	2.8
Ethnicity missing	0.0	0.0	0.0	0.6	0.0	0.6
Has some type of disability						
No	82.9	82.7	83.6	82.8	82.6	82.4
Yes	15.6	15.5	15.6	16.6	16.9	16.3
Missing disability status	1.6	1.8	0.9	0.6	0.5	1.2
Ever identified as gifted						
No	88.0	81.9	77.6	77.0	80.0	77.7
Yes	10.3	16.3	21.5	22.3	19.5	21.0
Missing gifted status	1.7	1.8	0.9	0.7	0.6	1.3
Over-age for grade						
No	68.4	69.4	76.3	77.3	77.6	76.4
Yes	31.5	29.3	23.1	22.1	21.9	22.5
Missing birthdate	0.0	1.3	0.6	0.5	0.5	1.1
Under-age for grade						
No	97.9	96.4	97.7	98.1	98.5	97.5
Yes	2.1	2.3	1.7	1.4	1.0	1.4
Missing birthdate	0.0	1.3	0.6	0.5	0.5	1.1
Ever repeated grade						
No	96.0	96.3	97.5	98.2	98.5	97.5
Yes	3.9	3.6	2.5	1.8	1.5	1.9
Missing repeater status	0.1	0.1	0.0	0.0	0.0	0.6

Administrative data from North Carolina provided by the NCERDC; An LEA is a Local Education Authority, akin to a school district.

Table 4.2 continued

Student and teacher characteristics across comparison samples

	TFA Students	Non-TFA, Same School-Grade-Year	Non-TFA, Same School, Different Grade/Yr	Non-TFA, Same LEA, Different School	Non-TFA, Different LEA	Total
Standardized Lagged Test Scores (gr 3-8)						
Math - Mean	-0.31	-0.20	-0.03	0.03	0.00	0.01
Math - SD	0.86	0.91	0.88	0.82	0.81	0.83
%Missing	58.34	66.87	44.46	41.56	42.12	43.66
Reading - Mean	-0.31	-0.21	-0.04	0.03	0.01	0.00
Reading - SD	0.87	0.92	0.87	0.81	0.82	0.83
%Missing	58.32	66.87	44.44	41.55	42.12	43.65
Standardized 8th Grade Test Scores (gr 9-12)						
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
%Missing	16.43	18.92	26.12	22.75	20.56	22.61
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
%Missing	16.59	19.09	26.24	22.71	20.64	22.64
Years of Experience						
Mean	0.9	11.4	12.2	12.8	15.1	13.0
SD	1.5	9.5	9.7	9.5	9.6	9.7
Yrs of experience missing	50.0	36.5	38.9	40.8	61.0	45.6
% Veteran Teacher (>3 Years experience)						
No	46.7	15.0	13.5	10.9	4.6	10.1
Yes	3.3	48.5	47.6	48.3	34.5	44.3
Missing teacher experience	50.0	36.5	38.9	40.8	61.0	45.6
Fully Certified						
No	55.4	7.9	4.3	3.3	2.7	3.7
Yes	26.5	77.6	74.6	79.7	76.5	77.2
Missing teacher cert.	18.1	14.4	21.1	17.0	20.8	19.1
Has an MA or more						
No	90.4	57.9	55.1	58.5	55.1	56.8
Yes	4.7	31.2	25.8	26.4	25.6	26.1
Missing degree status	4.9	10.9	19.1	15.1	19.3	17.2
Barron's Selectivity Rankings						
University Unranked	94.1	88.7	80.6	84.8	80.6	82.7
Most Competitive	0.0	0.1	0.1	0.0	0.0	0.0
Highly Competitive Plus	0.2	0.1	0.0	0.0	0.0	0.0
Highly Competitive	0.7	0.2	0.2	0.1	0.0	0.1
Very Competitive	0.0	0.0	0.0	0.0	0.0	0.0
University Unknown	4.9	10.9	19.1	15.1	19.3	17.2
Praxis Certification Score						
Mean	0.5	0.1	0.1	0.2	0.2	0.2
SD	0.7	0.9	0.9	0.9	0.8	0.9
Praxis score missing	21.0	30.8	31.3	26.4	27.2	28.1
N (Students)	116715.0	1150219.0	5427891.0	14753375.0	7046688.0	28494888.0
N (Teachers)	904	4,741	36,198	89,832	42,350	174,025

Administrative data from North Carolina provided by the NCERDC; An LEA is a Local Education Authority, akin to a school district.

Table 4.3

Impact of TFA on achievement in elementary, middle, and high school, school-grade-year or school-subject-year fixed effects

	Elementary Math	**Elementary Reading**	**Middle School Math**	**Middle School Reading**	**High School English**	**High School Math**	**High School Science**	**High School Social Studies**
TFA	**0.040*** **	-0.019	**0.046*** **	**0.023*** **	**0.036*** **	**0.112*** **	**0.197*** **	**0.050*** **
	(0.010)	(0.010)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.253***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.566***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score Missing	-0.343***	0.478***	-0.328***	0.514***	0.361***	-0.107***	0.137***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.198***	-0.186***	-0.283***	-0.280***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.209***	0.102***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.103***	-0.142***	-0.112***	-0.164***	-0.177***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.027***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.125***	0.167***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 4.3 continued

Impact of TFA on achievement in elementary, middle, and high school, school-grade-year or school-subject-year fixed effects

	Elementary Math	**Elementary Reading**	**Middle School Math**	**Middle School Reading**	**High School English**	**High School Math**	**High School Science**	**High School Social Studies**
Parent Ed: College or more	0.148***	0.154***	0.139***	0.159***	0.173***	0.158***	0.162***	0.193***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.327***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.417***	-0.090**	-0.063	-0.238***
	(.)	(.)	(.)	(.)	(0.032)	(0.030)	(0.033)	(0.029)
Student is gifted	0.712***	0.616***	0.400***	0.304***	0.326***	0.336***	0.357***	0.370***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.615***	-0.429***	-0.538***	-0.590***	-0.242**	-0.353***	-0.194	-0.287***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.146***	-0.246***	-0.320***	-0.362***	-0.505***	0.080	-0.218**	-0.301***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.170***	-0.168***	-0.129***	-0.121***	-0.202***	-0.189***	-0.237***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.635***	-0.630***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.280*	-0.823***	-0.461***	-0.604***	-0.249***	-0.409***	-0.672***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.075***	0.013***	-0.020***	-0.029***	-0.004*	-0.161***	-0.037***	-0.016***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.533	0.501	0.657	0.605	0.594	0.428	0.415	0.429
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 4.4

Impact of TFA on achievement, elementary, middle, and high school; controlling for teacher quality[1]

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
Praxis Scores								
TFA	**0.029****	**-0.023***	**0.041*****	**0.022*****	**0.026****	**0.093*****	**0.179*****	0.024
	(0.010)	(0.011)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Mean Praxis Score	**0.016*****	**0.004*****	**0.009*****	**0.003*****	**0.014*****	**0.013*****	**0.027*****	**0.033*****
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Years of Experience								
TFA	**0.051*****	-0.009	**0.051*****	**0.026*****	**0.046*****	**0.129*****	**0.215*****	**0.071*****
	(0.010)	(0.010)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Years of Experience	**0.001*****	**0.001*****	**0.001*****	**0.001*****	**0.002*****	**0.002*****	**0.002*****	**0.003*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Veteran teachers								
TFA	**0.058*****	-0.003	**0.057*****	**0.029*****	**0.048*****	**0.143*****	**0.229*****	**0.092*****
	(0.010)	(0.011)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Veteran Teacher (>3 years experience)	**0.038*****	**0.035*****	**0.026*****	**0.015*****	**0.053*****	**0.093*****	**0.090*****	**0.089*****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Fully Certified								
TFA	**0.074*****	0.004	**0.066*****	**0.038*****	**0.062*****	**0.177*****	**0.256*****	**0.115*****
	(0.010)	(0.011)	(0.006)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Fully Certified	**0.066*****	**0.045*****	**0.042*****	**0.032*****	**0.091*****	**0.152*****	**0.124*****	**0.151*****
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Has an MA or more								
TFA	**0.040*****	-0.019	**0.048*****	**0.023*****	**0.034*****	**0.106*****	**0.198*****	**0.052*****
	(0.010)	(0.011)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Has MA or more	**0.006*****	0.002	**0.010*****	0.001	**0.004***	**0.010*****	**0.022*****	**0.026*****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)

[1] Summary of TFA and teacher quality coeffficients. Full models shown in Appendix Tables 4.3 - 4.8.

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Models include fixed effects for school-grade-year or school-subject-year combinations.

Table 4.4 continued

Impact of TFA on achievement, elementary, middle, and high school; controlling for teacher quality[1]

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
				Barron's Rankings				
TFA	**0.038***** (0.010)	-0.020 (0.010)	**0.044***** (0.005)	**0.023***** (0.006)	**0.034***** (0.010)	**0.104***** (0.009)	**0.190***** (0.011)	**0.046***** (0.012)
Barron's - Most Competitive	**0.100***** (0.016)	**0.034*** (0.017)	**0.074***** (0.021)	0.031 (0.019)	-0.108 (0.143)	**-0.070**** (0.024)	**0.454***** (0.025)	**0.058*** (0.026)
Barron's - Highly Competitive Plus	-0.044 (0.057)	-0.073 (0.059)	-0.047 (0.029)	**-0.051*** (0.025)	0.020 (0.023)	**-0.186***** (0.052)	**0.093***** (0.028)	-0.003 (0.023)
Barron's - Highly Competitive	**0.037*** (0.014)	0.022 (0.015)	**0.042***** (0.009)	0.004 (0.010)	**-0.046*** (0.021)	-0.032 (0.016)	0.000 (0.022)	**-0.132***** (0.013)
Barron's - Very Competitive Plus	-0.056 (0.043)	-0.054 (0.041)	0.062** (0.019)	0.026 (0.025)	0.021 (0.046)	**-0.230**** (0.080)	**-0.143**** (0.050)	0.055 (0.041)

[1] Summary of TFA and teacher quality coeffficients. Full models shown in Appendix Tables 4.3 - 4.8.

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Models include fixed effects for school-grade-year or school-subject-year combinations.

Table 4.5

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

	Elementary Math			Elementary Reading			Middle School Math			Middle School Reading		
	2001-2011	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**
TFA	**0.040*****	0.008	**0.042*****	-0.019	**-0.086***	-0.015	**0.046*****	0.016	**0.052*****	**0.023*****	0.004	**0.027*****
	(0.010)	(0.038)	(0.011)	(0.010)	(0.040)	(0.011)	(0.005)	(0.012)	(0.006)	(0.006)	(0.013)	(0.006)
Lagged Reading Score	0.022***	-0.000	0.043***	0.451***	0.407***	0.487***	0.092***	0.058***	0.124***	0.529***	0.491***	0.566***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Reading Score is Missing	0.064***	0.092***	0.043***	-0.669***	-0.674***	-0.636***	0.002	0.053***	-0.042***	-0.784***	-0.781***	-0.747***
	(0.006)	(0.009)	(0.008)	(0.008)	(0.012)	(0.011)	(0.004)	(0.007)	(0.005)	(0.005)	(0.008)	(0.007)
Lagged Math Score	0.497***	0.476***	0.517***	0.070***	0.063***	0.081***	0.576***	0.560***	0.596***	0.132***	0.105***	0.162***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Math Score is Missing	-0.343***	-0.391***	-0.300***	0.478***	0.461***	0.471***	-0.328***	-0.375***	-0.276***	0.514***	0.503***	0.497***
	(0.006)	(0.009)	(0.008)	(0.008)	(0.012)	(0.011)	(0.004)	(0.007)	(0.006)	(0.005)	(0.008)	(0.007)
Black	-0.309***	-0.321***	-0.296***	-0.288***	-0.293***	-0.282***	-0.163***	-0.210***	-0.113***	-0.194***	-0.244***	-0.142***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.057***	-0.070***	-0.099***	-0.071***	-0.110***	-0.038***	-0.052***	-0.025***	-0.059***	-0.075***	-0.045***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
Asian	0.117***	0.081***	0.137***	-0.037***	-0.072***	-0.019***	0.160***	0.149***	0.174***	-0.018***	-0.031***	-0.004
	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)
American Indian	-0.151***	-0.156***	-0.147***	-0.165***	-0.159***	-0.169***	-0.088***	-0.106***	-0.068***	-0.103***	-0.121***	-0.082***
	(0.003)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)
Other Race	-0.105***	-0.103***	-0.107***	-0.085***	-0.073***	-0.093***	-0.048***	-0.063***	-0.034***	-0.031***	-0.050***	-0.015***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
Female	-0.070***	-0.068***	-0.072***	0.045***	0.048***	0.042***	-0.027***	-0.041***	-0.016***	0.042***	0.048***	0.035***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.133***	0.046***	0.117***	0.154***	0.058***	0.071***	0.128***	0.026***	0.101***	0.170***	0.044***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.214***	0.268***	0.087***	0.230***	0.290***	0.095***	0.147***	0.209***	0.080***	0.170***	0.247***	0.088***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: College or more	0.148***	0.455***	0.069***	0.154***	0.491***	0.071***	0.139***	0.175***	0.095***	0.159***	0.208***	0.104***
	(0.004)	(0.016)	(0.004)	(0.004)	(0.017)	(0.004)	(0.002)	(0.014)	(0.002)	(0.002)	(0.016)	(0.002)
Student has disability	-0.319***	-0.346***	-0.291***	-0.397***	-0.433***	-0.359***	-0.208***	-0.279***	-0.135***	-0.271***	-0.357***	-0.184***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student is gifted	0.712***	0.736***	0.680***	0.616***	0.630***	0.594***	0.400***	0.504***	0.288***	0.304***	0.387***	0.213***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

* p<0.05 ** p<0.01 *** p<0.001;

Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 4.5 continued

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

	Elementary Math			Elementary Reading			Middle School Math			Middle School Reading		
	2001-2011	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**
Gifted status missing	-0.615***	0.000	-0.619***	-0.429***	0.000	-0.430***	-0.538***	0.000	-0.533***	-0.590***	0.000	-0.590***
	(0.031)	(.)	(0.031)	(0.048)	(.)	(0.048)	(0.030)	(.)	(0.029)	(0.051)	(.)	(0.049)
Student ever IDed as LEP	-0.124***	-0.083***	-0.143***	-0.252***	-0.194***	-0.279***	-0.032***	-0.051***	-0.019***	-0.152***	-0.167***	-0.141***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
LEP status missing	-0.146***	0.000	-0.159***	-0.246***	0.000	-0.259***	-0.320***	1.817*	-0.335***	-0.362***	0.000	-0.380***
	(0.025)	(.)	(0.025)	(0.027)	(.)	(0.027)	(0.025)	(0.775)	(0.024)	(0.028)	(.)	(0.027)
Student over age for grade	-0.170***	-0.160***	-0.177***	-0.168***	-0.178***	-0.157***	-0.129***	-0.137***	-0.119***	-0.121***	-0.132***	-0.107***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.055***	0.068***	0.049***	0.041***	0.058***	0.065***	0.066***	0.061***	0.035***	0.035***	0.035***
	(0.004)	(0.005)	(0.006)	(0.004)	(0.006)	(0.006)	(0.003)	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)
Student has repeated grade	-0.892***	-0.867***	-0.914***	-0.978***	-0.993***	-0.944***	-0.421***	-0.423***	-0.396***	-0.467***	-0.491***	-0.404***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)
Repeater status missing	-0.280*	0.003	-0.584**	-0.823***	-0.798**	-0.845***	-0.461***	-0.289**	-0.568***	-0.604***	-0.243*	-0.837***
	(0.135)	(0.189)	(0.193)	(0.153)	(0.270)	(0.185)	(0.063)	(0.099)	(0.081)	(0.073)	(0.120)	(0.090)
Constant	0.075***	0.054***	0.133***	0.013***	-0.016***	0.079***	-0.020***	-0.034***	-0.017***	-0.029***	-0.060***	-0.010***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
R-squared	0.533	0.545	0.525	0.501	0.509	0.498	0.657	0.652	0.670	0.605	0.586	0.630
Observations	3793320	1829787	1963533	3770563	1819654	1950909	3734293	1822606	1911687	3722294	1817856	1904438

* p<0.05 ** p<0.01 *** p<0.001;

Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 4.6

Impact of TFA on same-year tests in high school with school-subject-year fixed effects, by early and established TFA

	High School English			High School Math			High School Science			High School Social Studies		
	2001-2011	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**
TFA	**0.036*****	0.020	**0.041*****	**0.112*****	**0.158*****	**0.107*****	**0.197*****	**0.126*****	**0.216*****	**0.050*****	**-0.350*****	**0.063*****
	(0.010)	(0.021)	(0.010)	(0.009)	(0.032)	(0.009)	(0.011)	(0.021)	(0.012)	(0.012)	(0.075)	(0.012)
8th grade reading score	0.440***	0.422***	0.456***	0.061***	0.044***	0.077***	0.253***	0.237***	0.277***	0.389***	0.370***	0.417***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
8th grade reading score is missing	-0.573***	-0.598***	-0.610***	0.002	0.144***	-0.027**	-0.088***	-0.060**	-0.119***	-0.212***	-0.226***	-0.237***
	(0.009)	(0.019)	(0.010)	(0.009)	(0.022)	(0.009)	(0.011)	(0.023)	(0.012)	(0.010)	(0.032)	(0.011)
8th grade math score	0.168***	0.159***	0.180***	0.566***	0.549***	0.582***	0.348***	0.313***	0.389***	0.186***	0.120***	0.231***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
8th grade math score is missing	0.361***	0.479***	0.308***	-0.107***	-0.246***	-0.068***	0.137***	0.143***	0.136***	0.341***	0.397***	0.357***
	(0.009)	(0.019)	(0.010)	(0.009)	(0.022)	(0.009)	(0.011)	(0.023)	(0.012)	(0.010)	(0.032)	(0.011)
Black	-0.198***	-0.233***	-0.162***	-0.186***	-0.238***	-0.135***	-0.283***	-0.351***	-0.199***	-0.280***	-0.396***	-0.188***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Hispanic	-0.085***	-0.131***	-0.064***	-0.044***	-0.057***	-0.033***	-0.084***	-0.126***	-0.054***	-0.059***	-0.124***	-0.041***
	(0.003)	(0.005)	(0.004)	(0.002)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.003)	(0.006)	(0.004)
Asian	-0.025***	-0.011	-0.029***	0.209***	0.179***	0.239***	0.102***	0.077***	0.142***	0.013***	-0.038***	0.054***
	(0.004)	(0.006)	(0.005)	(0.003)	(0.005)	(0.004)	(0.003)	(0.005)	(0.005)	(0.004)	(0.007)	(0.005)
American Indian	-0.142***	-0.163***	-0.115***	-0.112***	-0.126***	-0.096***	-0.164***	-0.190***	-0.125***	-0.177***	-0.207***	-0.143***
	(0.006)	(0.008)	(0.007)	(0.004)	(0.007)	(0.006)	(0.005)	(0.008)	(0.007)	(0.006)	(0.010)	(0.007)
Other Race	-0.028***	-0.030***	-0.018***	-0.055***	-0.043***	-0.051***	-0.057***	-0.056***	-0.040***	-0.068***	-0.064***	-0.048***
	(0.004)	(0.007)	(0.004)	(0.003)	(0.006)	(0.003)	(0.004)	(0.007)	(0.004)	(0.004)	(0.009)	(0.004)
Female	0.137***	0.141***	0.133***	-0.020***	-0.042***	-0.002	-0.142***	-0.168***	-0.114***	-0.202***	-0.198***	-0.205***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Parent Ed: HS or less	0.080***	0.138***	0.040***	0.036***	0.072***	0.002	0.082***	0.133***	0.017***	0.111***	0.197***	0.030***
	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Parent Ed: Some college/trade school	0.139***	0.197***	0.096***	0.080***	0.125***	0.034***	0.125***	0.189***	0.040***	0.167***	0.272***	0.074***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Parent Ed: College or more	0.173***	-0.716***	0.184***	0.158***	-0.328***	0.117***	0.162***	-0.394***	0.116***	0.193***	-0.192***	0.124***
	(0.003)	(0.029)	(0.003)	(0.002)	(0.025)	(0.003)	(0.003)	(0.027)	(0.003)	(0.003)	(0.027)	(0.003)
Student has disability	-0.327***	-0.400***	-0.268***	-0.133***	-0.181***	-0.103***	-0.165***	-0.257***	-0.088***	-0.165***	-0.386***	-0.055***
	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Disability status missing	-0.417***	0.000	-0.383***	-0.090**	0.000	-0.104***	-0.063	0.000	-0.073*	-0.238***	0.353	-0.254***
	(0.032)	(.)	(0.031)	(0.030)	(.)	(0.029)	(0.033)	(.)	(0.032)	(0.029)	(0.539)	(0.028)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 4.6 continued

Impact of TFA on same-year tests in high school with school-subject-year fixed effects, by early and established TFA

	High School English			High School Math			High School Science			High School Social Studies		
	2001-2011	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**
Student is gifted	0.326***	0.400***	0.256***	0.336***	0.424***	0.253***	0.357***	0.472***	0.211***	0.370***	0.588***	0.227***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Gifted status missing	-0.242**	0.000	-0.231**	-0.353***	0.000	-0.333***	-0.194	0.000	-0.166	-0.287***	0.000	-0.269***
	(0.083)	(.)	(0.080)	(0.057)	(.)	(0.056)	(0.099)	(.)	(0.095)	(0.072)	(.)	(0.069)
Student ever IDed as LEP	-0.223***	-0.263***	-0.208***	-0.001	-0.002	-0.005	-0.116***	-0.163***	-0.103***	-0.163***	-0.322***	-0.135***
	(0.003)	(0.005)	(0.004)	(0.002)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.003)	(0.007)	(0.004)
LEP status missing	-0.505***	0.000	-0.480***	0.080	0.532	0.055	-0.218**	0.000	-0.239***	-0.301***	0.000	-0.316***
	(0.060)	(.)	(0.057)	(0.046)	(0.473)	(0.045)	(0.069)	(.)	(0.066)	(0.053)	(.)	(0.051)
Student over age for grade	-0.202***	-0.222***	-0.188***	-0.189***	-0.200***	-0.180***	-0.237***	-0.257***	-0.215***	-0.244***	-0.292***	-0.209***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Student under age for grade	0.109***	0.110***	0.102***	0.151***	0.167***	0.132***	0.190***	0.207***	0.171***	0.165***	0.172***	0.160***
	(0.005)	(0.008)	(0.007)	(0.004)	(0.005)	(0.005)	(0.004)	(0.006)	(0.006)	(0.004)	(0.008)	(0.005)
Student has repeated grade	-0.421***	-0.544***	-0.359***	-0.513***	-0.572***	-0.472***	-0.635***	-0.670***	-0.593***	-0.630***	-0.777***	-0.552***
	(0.002)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)	(0.007)	(0.006)	(0.005)	(0.009)	(0.005)
Repeater status missing	-0.249***	-0.529***	-0.218***	-0.409***	-0.679***	-0.373***	-0.672***	-0.727***	-0.665***	-0.477***	-0.941***	-0.448***
	(0.022)	(0.073)	(0.023)	(0.021)	(0.063)	(0.022)	(0.032)	(0.069)	(0.034)	(0.028)	(0.107)	(0.028)
Constant	-0.004*	-0.078***	0.034***	-0.161***	-0.170***	-0.152***	-0.037***	-0.106***	0.032***	-0.016***	-0.098***	0.039***
	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
R-squared	0.594	0.571	0.621	0.428	0.401	0.457	0.415	0.387	0.459	0.429	0.382	0.476
Observations	1270188	591981	678207	2622518	1233805	1388713	2174625	1133522	1041103	1865239	680396	1184843

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 4.7

Impact of TFA in elementary, middle, and high school; testing for differences in early and later TFA

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	0.009	**-0.087***	0.016	0.006	0.012	**0.158*****	**0.130*****	**-0.473*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.034	0.073	**0.038****	0.022	0.031	-0.050	**0.091*****	**0.539*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.253***	0.389***
	(0.001)	(0.001)	0.000	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Reading Score is Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.566***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Math Score is Missing	-0.343***	0.478***	-0.328***	0.514***	0.361***	-0.107***	0.138***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.198***	-0.186***	-0.283***	-0.280***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.209***	0.102***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.103***	-0.142***	-0.112***	-0.164***	-0.177***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.027***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.100***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.125***	0.167***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Missing	0.148***	0.154***	0.139***	0.159***	0.173***	0.158***	0.162***	0.193***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.327***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status is missing	0.000	0.000	0.000	0.000	-0.417***	-0.090**	-0.063	-0.238***
	(.)	(.)	(.)	(.)	(0.032)	(0.030)	(0.033)	(0.029)
Student is gifted	0.712***	0.616***	0.400***	0.304***	0.326***	0.336***	0.357***	0.370***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.615***	-0.429***	-0.538***	-0.590***	-0.242**	-0.353***	-0.194	-0.287***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.146***	-0.246***	-0.320***	-0.362***	-0.505***	0.080	-0.218**	-0.301***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.170***	-0.168***	-0.129***	-0.121***	-0.202***	-0.189***	-0.237***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.630***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.280*	-0.823***	-0.461***	-0.604***	-0.249***	-0.409***	-0.672***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.075***	0.013***	-0.020***	-0.029***	-0.004*	-0.161***	-0.037***	-0.016***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.533	0.501	0.657	0.605	0.594	0.428	0.415	0.429
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Table 4.8

Impact of TFA in elementary, middle, and high school; using school-grade-year fixed effect;, testing whether teacher quality can explain time-trends in TFA effectiveness[1]

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
				Praxis Scores				
TFA	0.015	**-0.082***	0.019	0.009	0.012	**0.176*****	**0.130*****	**-0.440*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.014	0.064	**0.028***	0.017	0.020	**-0.088****	**0.071****	**0.482*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Mean Praxis Score	0.014***	0.004***	0.008***	0.004***	0.022***	0.019***	0.034***	0.043***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean Praxis	**0.004*****	-0.001	**0.003*****	**-0.002***	**-0.015*****	**-0.012*****	**-0.017*****	**-0.017*****
Score*Post-2005	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
				Years of Experience				
TFA	0.020	-0.077	0.019	0.009	0.032	**0.173*****	**0.167*****	**-0.423*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.032	0.072	**0.041****	0.021	0.016	-0.052	**0.060***	**0.502*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Years of	0.002***	0.001***	0.001***	0.001***	0.003***	0.003***	0.003***	0.005***
Experience	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of	**-0.001*****	**-0.000***	-0.000	**-0.001*****	**-0.002*****	**-0.002*****	**-0.003*****	**-0.003*****
Experience*Post-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
				Veteran Teacher				
TFA	0.027	-0.072	**0.026***	0.011	0.034	**0.174*****	**0.173*****	**-0.392*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.030	0.072	**0.040****	0.020	0.014	-0.041	**0.069****	**0.488*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Veteran teacher	0.045***	0.038***	0.028***	0.020***	0.078***	0.115***	0.116***	0.123***
(>3 yrs. expr.)	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
Veteran teacher*	**-0.015*****	**-0.006****	**-0.003***	**-0.013*****	**-0.049*****	**-0.042*****	**-0.054*****	**-0.057*****
Post-2005	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
				Fully Certified				
TFA	0.053	-0.055	**0.042*****	**0.026***	**0.075*****	**0.233*****	**0.211*****	**-0.339*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.015	0.055	**0.030***	0.015	-0.018	**-0.068***	**0.060***	**0.470*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Fully Certified	0.074***	0.053***	0.045***	0.034***	0.110***	0.175***	0.132***	0.146***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
Fully Certified*	**-0.025*****	**-0.023*****	**-0.007*****	**-0.005***	**-0.035*****	**-0.042*****	**-0.015*****	**0.007***
Post-2005	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
				Has an MA or More				
TFA	0.003	**-0.090***	0.016	0.005	0.007	**0.133*****	**0.122*****	**-0.475*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.040	0.074	**0.042****	0.022	0.035	-0.030	**0.100*****	**0.542*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Has MA or more	0.008***	0.006***	0.008***	0.005***	0.011***	0.019***	0.035***	0.034***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
Has MA or more*	**-0.004***	**-0.007*****	**0.005****	**-0.008*****	**-0.013*****	**-0.015*****	**-0.023*****	**-0.013*****
Post-2005	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)

[1] Summary of TFA and teacher quality coeffficients. Full models shown in Appendix Tables 4.4 (a-f).

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Models include fixed effects for school-grade-year or school-subject-year combinations.

Table 4.8 continued

Impact of TFA in elementary, middle, and high school; using school-grade-year fixed effect;, testing whether teacher quality can explain time-trends in TFA effectiveness[1]

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
				Barron's Selectivity Rankings				
TFA	0.001	**-0.092***	0.014	0.005	0.005	**0.129*****	**0.122*****	**-0.480*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.040	0.077	**0.039****	0.023	0.037	-0.027	**0.095*****	**0.542*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Most Competitive	0.088***	0.044*	0.072**	0.039	-0.108	-0.073*	0.494***	-0.007
	(0.020)	(0.021)	(0.023)	(0.025)	(0.143)	(0.033)	(0.029)	(0.043)
Highly	-1.665**	-0.073	-0.076	-0.029	0.205***	-0.470***	0.151***	-0.213**
Competitive Plus	(0.633)	(0.059)	(0.041)	(0.045)	(0.041)	(0.096)	(0.039)	(0.069)
Highly	0.009	0.038	-0.002	-0.019	0.003	-0.024	-0.179***	-0.027
Competitive	(0.028)	(0.029)	(0.016)	(0.017)	(0.033)	(0.023)	(0.033)	(0.029)
Very Competitive	-0.081	-0.099	0.043	0.007	0.021	-0.230**	-0.140	0.056
Plus	(0.055)	(0.059)	(0.025)	(0.027)	(0.046)	(0.080)	(0.097)	(0.041)
Most Competitive	0.035	-0.026	0.013	-0.020	0.000	0.007	**-0.165****	0.102
***Post-2005**	(0.034)	(0.034)	(0.060)	(0.039)	(.)	(0.049)	(0.058)	(0.053)
Highly Comp. Plus	**1.635***	0.000	0.058	-0.031	**-0.277*****	**0.403*****	**-0.122***	**0.236****
***Post-2005**	(0.635)	(.)	(0.058)	(0.054)	(0.050)	(0.115)	(0.055)	(0.073)
Highly Comp.	0.039	-0.021	**0.063*****	0.036	-0.079	-0.016	**0.339*****	**-0.134*****
***Post-2005**	(0.033)	(0.034)	(0.019)	(0.022)	(0.042)	(0.033)	(0.045)	(0.032)
Very Comp. Plus	0.061	0.089	0.050	0.120	0.000	0.000	0.006	0.000
***Post-2005**	(0.086)	(0.082)	(0.040)	(0.067)	(.)	(.)	(0.113)	(.)

[1] Summary of TFA and teacher quality coeffficients. Full models shown in Appendix Tables 4.4 (a-f).

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses. Models include fixed effects for school-grade-year or school-subject-year combinations.

Chapter 5

Summary and Conclusion

Teach For America (TFA) is an increasingly prominent part of the national dialogue on teacher training and recruitment. With teachers in urban and rural low-income school districts in 36 states, TFA has expanded across the country. While a growing body of research examines TFA's effects on student achievement, little of this work considers the ways in which TFA's effects on student outcomes might vary. I argue that it is important to consider variation in TFA's effectiveness, particularly since it seems unlikely to have uniform effects on all students given its wide-ranging presence. By understanding who benefits and suffers from having a TFA teacher (and how much), we can deepen our understanding of the costs and benefits of this controversial program, as well as the strengths and weaknesses of teacher training in the United States more broadly.

Summary of Findings

This dissertation examines variation in the effects of TFA across several dimensions, and considers different ways in which the fit between the TFA program and student and teacher characteristics impact student test scores. Chapter 2 examines whether TFA's effects vary across the distribution of elementary reading and math achievement using experimental data from Mathematica. In this chapter, I estimate quantile treatment effects (QTE) to examine TFA's effects on the post-test distributions of student math and reading achievement. I find that TFA has, on average, positive effects on elementary math that are shared across the distribution, which are as large as 0.34 standard deviations at some points of the distribution. In contrast, in reading, the average null effects found in previous research (Glazerman et al., 2006) are

concealing offsetting effects at different portions of the distribution. TFA has small positive effects on the upper portion of the distribution, with negative, marginally significant, effects on the bottom portion of the distribution. This pattern of results is even more pronounced when comparing TFA teachers with veteran non-TFA teachers, which yields statistically significant differences across most of the distribution, with TFA students doing worse at the bottom of the distribution (effect size = -0.19), and better at the top of the distribution (effect size = 0.19) .

Chapter 3 draws on theoretical and experimental work in developmental psychology to examine the stage-proficiency-environment fit of TFA. In other words, it examines whether the effects of TFA vary based on the developmental stage and academic skill level of the student to determine if there is a "best" match between student readiness and the timing of having a TFA teacher. This chapter uses administrative data from North Carolina from 2006-2011 to examine the relationship between having a TFA teacher and end of year student achievement test scores, estimating models with school-grade-year (or school-subject-year) fixed effects to compare students in TFA classrooms with students in non-TFA classrooms taking the same subject at the same school during the same school year. Additional models use student (or student-year) fixed effects that compare students with themselves across grade levels or across high school subjects where they did and did not have a TFA teacher. The results suggest that TFA has positive effects in nearly every subject and grade level, but that effects are largest in high school, particularly in math and science.

Chapter 3 also finds two distinct patterns relating to the proficiency-fit hypotheses, which vary by students' developmental stage. In elementary and middle school, TFA's effects adhere to the *accumulated advantages* hypothesis, where they are largest for students who enter with the highest skill levels. In contrast, in high school, TFA's effects follow a *compensatory* pattern,

where they are largest for students with the lowest entering skill levels. These results echo some of the findings from Chapter 2, as TFA teachers in North Carolina also appear to struggle with the lowest-performing readers. However, these results also extend our understanding of TFA by highlighting the particularly large beneficial effects of TFA in high school STEM subjects, with encouraging results that suggest TFA teachers are especially good at working with lower-skilled high school students in these subjects.

Chapter 4 examines whether TFA's effects have changed over time and whether the main effects of TFA, or changes in its effectiveness, can be explained by observable measures of teacher quality. This chapter also uses administrative data from North Carolina, but includes a longer panel that stretches from 2000-2011. Like Chapter 3, it uses OLS regression with school-grade-year and school-subject-year effects to estimate the quasi-experimental impact of TFA relative to teachers in the same schools, grades (or subjects), and years only. I find that TFA's effects have improved over time in three subject areas: middle school math, high school science, and high school social studies.

Most observable measures of teacher quality do not help account for the relationship between TFA and student achievement or the improvements in this relationship over time in some subjects. The one exception is average teacher Praxis scores, on which TFA teachers out-scored counterfactual teachers by nearly half a standard deviation. Praxis scores explain between a tenth and a third of the main effect of TFA and the change over time in TFA effectiveness, depending on the subject. Importantly, I find a TFA advantage of similar magnitude even when TFA teachers are compared exclusively with veteran teachers. The positive main effects of TFA in most subjects and the changes in their effectiveness over time in three subjects were consistent when comparing TFA teachers to veterans. The one notable change is that students in TFA

classrooms performed significantly worse than those in veteran counterfactual classrooms in elementary reading. These results again highlight TFA's positive effects in most subjects (except elementary reading), and suggest that the program's relative quality is improving over time. They also suggest that elements of teacher quality or TFA training that are not observable in the North Carolina data are largely responsible for TFA's effectiveness and improvements.

Key Lessons

Together, the three empirical chapters in this dissertation highlight several important findings about the TFA program and its effects on student achievement, many of which are consistent across chapters. First, TFA's largest and most consistent effects are in STEM subjects. TFA has positive effects on student math achievement in every grade. These effects are largest in high school, and for math range from an average effect size of 0.05 in elementary school to 0.20 in high school. Chapter 2 shows that these average effects may not capture the whole story, so that while TFA teachers raise the entire distribution of students in elementary math, at some points in the distribution the gap between TFA and non-TFA students is as large as 0.34 standard deviations. TFA also appears to be particularly effective with higher-performing students in grades 3-8, but has *compensatory* effects for high school students.

Although TFA teachers are good at improving math scores, they are even better at teaching science (relative to counterfactual teachers). Further, TFA teachers have been more effective in science than counterfactual teachers since at least the early 2000s, and have become relatively more effective at science instruction over time. This suggests that administrators and traditional teacher training programs would do well to improve their training and selection in math and science, as there would appear to be substantial room for improvement. Additionally, administrators should not be concerned about recruiting TFA teachers for science and math

positions, particularly in high school. Further, they should not worry about giving TFA teachers students from weaker math backgrounds in high school science or math, as TFA teachers seem to be especially effective at working with such students.

Second, TFA teachers struggle to teach elementary reading on average, and especially struggle with low-performing readers relative to non-TFA teachers. Chapter 2 showed that TFA teachers had a null or negative impact on the bottom of the distribution, and Chapter 3 suggested that TFA teachers are particularly ill-equipped to support students who enter well below grade level. In both cases, the negative impacts were larger when compared with veteran teachers. Elementary reading is one area in which the TFA program continues to struggle, although the analyses examining temporal changes suggest that TFA may have improved over time from being worse than counterfactual teachers to being equally effective. This also suggests that the challenges presented by working with low-performing early readers cannot be addressed by teacher selectivity alone. One might speculate, for example, that many of these students may have (or will develop) learning disabilities, and that TFA training does not equip teachers particularly well to address this. When it comes to early reading, TFA teachers are missing something that comes with experience or a more traditional certification pathway. In particular, TFA teachers may be missing early literacy courses that devote a substantial amount of course time to topics such as foundations of print, phonemic awareness, and phonological awareness, all of which are part of courses typically included in many elementary education and liberal studies undergraduate majors, as well as many traditional credential programs.

Third, excluding elementary reading, TFA teachers are actually better than veteran teachers in most subjects and grades. These findings are echoed in both chapters that compare TFA teachers with veteran non-TFA teachers. To be sure, TFA teachers perform particularly

poorly compared with veteran teachers in elementary reading[38], but in most other subjects, positive TFA effects are similar when compared with teachers in the same schools, grades (or subjects), and school years, regardless of the seniority of the counterfactual teacher. It is also important to note that while the teacher quality measures did little to explain the effectiveness of TFA, counterfactual teachers in these analyses are of relatively poor quality on the observable dimensions included in the administrative data relative to teachers in non-TFA schools and districts, underscoring the challenge of staffing these positions with high-quality teachers (Boyd et al., 2005b; Clotfelter et al., 2006). Nonetheless, the positive TFA findings for most subjects, even relative to veteran teachers, suggest that it would be helpful to examine TFA's training program and its selection criteria in greater detail.

To provide an additional perspective on the magnitude of TFA's effects on student achievement, the return to having a veteran (4 years of experience or more) versus a novice teacher ranges from 0.02 to 0.09 standard deviations. Relative to the corresponding effect sizes of TFA's impacts on student achievement across the corresponding subjects, TFA's effects are approximately twice as large as the effect of having a veteran relative to a novice teacher (with the exception of elementary reading). This suggests that having a TFA teacher may be more effective than improving efforts to retain veteran conventional teachers. However, given the documented costs of the turnover that results from employing TFA teachers, the relative benefits of having a TFA teacher may be somewhat offset (Boyd et al., 2006; Kane et al., 2008). While this issue has been investigated elsewhere in elementary and middle school, it has not yet been evaluated in the context of North Carolina or in subject areas where TFA teachers are especially effective relative to counterfactual teachers, such as high school science. A further investigation how the effectiveness of TFA may be offset by higher levels of turnover in high school subjects

would also help administrators and policy makers to weigh the relative costs and benefits of TFA teachers.

Fourth, although TFA's effects are largely positive, they still have a long way to go to help their students catch up with their more affluent peers. The effects identified across the three studies fall well within the range of effects of other well-regarded educational interventions. The magnitude of the TFA results across the studies are similar to the .22 SD effect observed in the Tennessee STAR class size experiment for early grades (Krueger, 1999), and the 0.35 SD effect of KIPP schools (Angrist et al., 2010). At the same time, students in TFA classrooms in North Carolina enter classrooms with scores that are 0.3 standard deviations below the state average and 0.4 standard deviations below students in low-poverty schools. In high school science, students with TFA teachers have end of course scores that are, on average, 0.20 standard deviations higher than those of students in non-TFA classrooms, and in math this difference is 0.11 standard deviations. Even if this effect were shared across the distribution and consistent for all students, it would not be enough to help students catch up to the average score of their peers across the state.[39]

To get a sense of the magnitude of the effects of TFA, it is potentially helpful to compare them with the average annual gains in effect sizes calculated by Bloom, Hill, Black, and Lipsey (2011) from a variety of nationally normed assessments. This backdrop reveals that several of the effects identified in this dissertation are substantial. The positive effect of TFA in high school science of 0.20 standard deviations is similar in magnitude to the average annual effect size gains of 0.15 to 0.19 standard deviations identified for high school science between grades 9 through 11. While Bloom and colleagues' calculations were taken from a group of nationally normed tests that do not include the end of course and end of grade tests used in North Carolina, and they

are grade specific rather than subject specific, the relative benchmark they provide suggests that having a TFA teacher for science in high school corresponds to an additional year of gains in high school relative to having a counterfactual teacher from the same school. For high school math, the same comparison suggests that the TFA effects identified in North Carolina of 0.11 standard deviations correspond to an additional half to three-quarters of a year of math instruction. While these are substantial gains, they still fall short of the 1.5 to 2 years of growth aspired to by TFA in its stated goals. Potentially more problematic than not achieving an ambitious goal, however, are the findings for TFA in elementary reading, where TFA has marginally significant negative effects (ES = -0.19). Surprisingly, however, the magnitude of this effect corresponds to a reduction of only five percent of the annual gains identified for elementary reading by Bloom and colleagues' annual gains metric.

This metric can also help us understand the magnitude of the effects for the changes over time in the TFA effects reported in Chapter 4. Using Bloom and colleague's average annual gains reference for the gains in middle school math suggests that TFA effectiveness has increased by ten percent of the usual yearly gains relative to the earlier time period. In high school science, the magnitude of the change over time corresponds to an improvement of approximately half of a school year. In high school social studies, this change is even larger, but should be interpreted with caution given the large negative effects of TFA in the earlier time period that are driving the results.

Considered in light of the typical gain achieved by students in different subjects and years, the effects of TFA are quite remarkable, as are the changes over time in these effects. However, at most, TFA teachers yield an additional year's worth of growth above what is produced by a counterfactual teacher, and effects of this magnitude are found only in subjects in

which their relative effectiveness is high, and where average annual gains in effect sizes are usually relatively small, such as high school science. Thus, while TFA is helping students make great strides in many subjects, and in some cases especially helps struggling students, these results suggest that TFA, as well as other traditional and alternative certification pathways and continuing teacher education programs, needs to continue improving in order to sufficiently raise teacher quality in underserved schools.

Fifth, all three studies highlight important aspects of variation in program effects that other researchers should consider when evaluating TFA and other interventions in the future. Analyses from each of the chapters identified different sources of variation in TFA's effects stemming from differences in the fit between the student the TFA program. Chapter 2 highlighted how examining distributions can uncover off-setting effects that appear to be null effects on average. Chapter 3 pointed to the ways in which the fit between the program and the child's readiness and developmental stage can interact with program effectiveness to provide additional gains or (dampen program impacts) for particular types of students. Chapter 4 underscored the need to consider the evolution of a program over time and the ways that this might intensify or dilute program impacts. All three chapters underscore one of the primary conclusions of this dissertation, documenting that while largely positive, TFA's effects are not uniform. This is true across the distribution, across subject areas, across developmental stages, across levels of student preparedness, and across chronological time.

Finally, this dissertation was guided by several literatures from developmental science and education that yielded theory-based expectations for heterogeneity in the effects of TFA. Given the variety of competing paradigms with countervailing predictions, it is not surprising that my findings were not congruent with all of these perspectives. Below I briefly review the

major paradigms that I drew upon, and offer some reflections on the implications of my results for these perspectives.

One major vein of research in developmental psychology that I drew on emphasizes the importance of the fit between the school environment and the developmental timing of the students (Eccles & Midgley, 1989). This theory predicts that, if the TFA intervention is effective, its impacts should be greatest at ages where the mismatch between students' needs and the school setting is the greatest, which occurs for most students in middle school.

Scholars in education and developmental psychology also highlight the importance of school structure and an integrated curriculum for supporting student learning and enhancing the impacts of interventions (Bredekamp & Rosegrant, 1992; Cobb, 1994; Piaget & Inhelder, 1969). This theoretical tradition suggests that TFA effects should be largest in elementary school, where TFA teachers can best support their students academically and emotionally.

Work from education also underscores the importance of teacher content knowledge for student learning (Hill et al., 2005). This work predicts that TFA's impacts should be largest in the areas where the corps members' content knowledge is strongest relative to that of counterfactual teachers, which is most likely in math and science, because the majority of traditionally trained teachers have far fewer undergraduate courses in these subjects and test score gaps are especially large in these subjects. In addition, this work suggests that TFA teachers should be more effective in later grades, where specific, technical content knowledge becomes more important.

Finally, work from developmental science, policy, and sociology yields conflicting theoretical predictions about whether TFA's effects will be largest for students who enter with

low, medium, or high skills (Cunha & Heckman, 2007; E. B. Miller et al., 2014; Rutter, 1987; Sameroff & Chandler, 1975).

My results examining developmental timing found that the effects of TFA were smallest in elementary school, and were sometimes negative in elementary reading. While the positive effect in math could be interpreted as the benefit of having an integrated, emotionally supportive classroom environment, it is unclear how to reconcile this with negative effects in reading. Likewise, while TFA middle school teachers do outperform their non-TFA counterparts in both math and reading, TFA's effects are largest in high school. This might suggest that misalignment of middle school structures and middle school students' developmental needs was so great that students were unable to fully take advantage of their TFA teachers. However, while in middle school we do not observe the stark differences between TFA math and reading impacts that I found in elementary school, important differences persist, with TFA teachers boosting math achievement more than reading achievement, and this finding suggests that there were other factors that overshadowed the stage-environment fit considerations.

On balance, the pattern of effects observed can best be explained by teacher content knowledge. While TFA teachers were effective in some elementary and middle school subjects, their effectiveness appears to be driven more by their relative strength in content-heavy courses in high school. In particular, TFA teachers outshine counterfactual teachers in STEM-related fields. Across all three school levels, TFA's effects were larger in math, and particularly science, than in ELA or social studies. This suggests that teacher content knowledge plays an important role in promoting TFA teacher effectiveness and that content knowledge outweighs the benefits of having a supportive teacher during a developmentally challenging period. The teacher content knowledge perspective is also consistent with finding somewhat larger effects in middle school

than elementary school, and can help explain the negative effects in elementary school reading. That is, we would not expect TFA teachers to have the specialized knowledge needed for interventions with struggling readers with things like phonemic awareness, and would rather expect traditionally certified teachers, particularly veterans, to have stronger skills and knowledge about how to navigate this. My results suggest that standard models conceptualizing the fit of an intervention might also consider differences across subjects, either as a robustness check (in case where no differences would be expected), or as an additional dimension along which a person and an intervention might "fit."

One set of findings where I found surprisingly little variation by subject, and clear differences by developmental stage, was in examining baseline skill interactions. Here I found evidence that at different developmental stages TFA functions in accordance with both the *accumulated advantages* and *compensatory* perspectives. In high school, where we might expect TFA teachers' superior content knowledge to enable them to particularly help students entering with strong skills, I found instead that TFA teachers had their largest gains with students who entered with low skills. This is in line with the *compensatory* hypothesis. By contrast, in elementary and middle school, I found evidence that TFA teachers particularly helped students who entered with high skills, in keeping with the *accumulated advantages* perspective. This developmental difference is intriguing, as one could imagine that strong content knowledge could be useful for helping challenge the students who came in with strong skills, and also that they might be helpful for intervening with struggling students who might need concepts explained to them in multiple ways. However, it is less clear from existing theory why we would expect to find different effects at different developmental stages, suggesting a fruitful arena for future theorizing.

Further, while the analyses and results in Chapter 4 were not conceptualized in terms of fit, viewed from this perspective they yield the intriguing insight that the relative fit between TFA and its students is improving over time in some subjects. While organizations targeting youth for interventions might strive to improve their fit with students as their program matures, the idea that a person-environment fit can evolve over time is not something that the original theory addressed, and could be a potentially helpful extension of this theory.

Future Research

This dissertation also points to several future avenues for research that would help to further deepen our understanding of TFA's effects. My future research priorities on TFA emerge from the need to replicate the analyses in this dissertation as well as explore additional questions that I was unable to answer with the experimental and administrative data I had available. First, it would be useful to replicate the distributional analyses from the experimental analysis in Chapter 2 using the administrative data from North Carolina and using the recently released Mathematica data with an updated experiment comparing TFA and non-TFA teachers in secondary math (Clark et al., 2013). In addition to replicating the elementary analyses, this work could be extended by examining how TFA impacts the distribution of student achievement in middle and high school. This would be particularly interesting given that high school math and science have large average impacts. It would also be interesting to compare the impacts of TFA on the post-test distribution with the pre-test interactions from Chapter 3 to examine mobility in the rank ordering of students during the course of a school year.

This dissertation also raises questions about the lasting impact of TFA. Research about teacher effectiveness more generally yields mixed conclusions about the persistence of teacher effects on test scores and other measures of attainment (Chetty, Friedman, & Rockoff, 2011;

Rothstein, 2010). Given that TFA's mission statement indicates an interest in creating long-lasting changes, it would also be of interest to the organization to learn whether TFA has any lasting impact on the students it serves. Future work might test whether TFA's impacts on test scores persist into the year after a student has a TFA teacher, and could also examine whether having a TFA teacher impacts several end of high school outcomes, including graduating, cumulative high school GPA, and planning to attend a two or four year college.

Thinking about the persistence of TFA impacts also raises interesting questions about whether TFA's effects are intensified by repeated exposure to TFA teachers. Here it would be interesting to consider not only whether the impacts of having multiple TFA teachers in the same subject build upon one another, but also what happens when a student has more than one TFA teacher across subjects within the same year. Such analyses might examine impacts on same-year test scores as well as end of high school outcomes.

The results from this dissertation also raise questions about the mechanisms behind TFA's relative successes in most subjects, particularly in high school math and science, as well as their relative struggles with teaching elementary reading, particularly with struggling readers. Observable teacher characteristics readily available in administrative data were unable to explain most of the relationship between TFA and student achievement, and thus additional data are needed. While many potential types of data might be of interest, two areas appear to be particularly promising for understanding why TFA teachers outperform their counterparts in most subjects, and why they struggle in elementary reading. First, it would be useful to learn more about the TFA recruitment and training process, to the extent that the organization would permit such an evaluation. In particular, observing the selection and training process, contrasting how TFA recruits are prepared to teach different types of subjects, with specific attention paid to

interventions for struggling students might yield important insights. An investigation of TFA's training and selection processes would be further strengthened by examinations of TFA's training materials over time, ideally contrasting them with materials from several well-regarded traditional teacher preparation programs. While this material may be difficult to obtain, it would help to deepen our understanding of what TFA is doing to prepare its corps members and which aspects of this preparation seem most beneficial for student learning.

In addition to learning more about TFA training, it would be helpful to compare teaching and planning practices between TFA and non-TFA teachers in the same schools using classroom observations and teacher interviews. Given the findings in this dissertation, observations of how TFA and non-TFA teachers differ in their approaches to supporting struggling learners, in addition to more general differences in instructional techniques and classroom management, would be potentially very informative.

These results also point to the need for continued replication across new data sets, beyond what I have proposed above, ideally across multiple TFA sites in a variety of states. This dissertation uses the best available experimental and administrative data to examine TFA program effects, and it confirms many of the previous findings identified by other research. However, it will be especially important to attempt to replicate the patterns of impact variation seen here in other settings to increase our confidence in these results and their policy implications.

Implications for Teacher Training

Finally, this dissertation also has implications for traditional teacher recruitment and preparation pathways, although the future research suggested above will further inform our understanding of how we can best move forward. My results contribute to the debate about

improving teacher quality and lend support to those who call for improving the quality of the types of individuals that become teachers, particularly among individuals who teach in low-income schools. Given that TFA teachers outperformed their same-school counterparts in most subjects and grade levels and that observable measures of teacher quality did not explain this relationship, other observable (or unobservable) characteristics identified by TFA's selection process are likely to be driving these results. It seems unlikely that TFA will be willing to disclose more data on its selection criteria than it shared with Dobbie (2011), but working to improve the quality of new teacher applicants in terms of college selectivity, undergraduate class rank, and standardized test performance might be a useful avenue for traditional teacher certification programs to pursue. Of course, much of this depends on making teaching in these schools more attractive, presumably by either improving the esteem and working conditions of teaching in low-income schools or by providing other types of incentives to accomplished teachers, which has proven difficult in the past (Glazerman et al., 2013).

However, while selectivity is no doubt important, recent evidence suggests that the TFA effect is unlikely to be attributable to selection alone, and that TFA's training also needs to be considered. The results from the most recent Mathematica random assignment study compared both TFA and Teaching Fellows teachers to counterfactual teachers in secondary math, showing that while TFA teachers were more effective than counterfactual teachers, the Teaching Fellows teachers were not (Clark et al., 2013). Teaching Fellows programs are also very selective, although perhaps not as selective as TFA. While this might suggest the need to further increase the selectivity of Teaching Fellows programs, it seems likely that some aspect of the TFA program is also contributing to its corps members' success that is lacking from the Teaching Fellows training program. Hopefully, future investigations of TFA training will yield some

insights about whether any particular aspects of TFA training can inform and improve other teacher preparation pathways.

Even if particular aspects of TFA recruitment and training could be translated to improve traditional teacher preparation programs, my results suggest that both types of training need further improvements to help low-income students catch up to their higher-income peers. This dissertation cannot pinpoint the particular aspects of teacher preparation that would result in substantial improvements. However, it does underscore the idea that, with a focused, selective program, some sizable improvements can be made. TFA is a small, but effective tool for improving achievement for some students, but system-level changes are needed to more fully address income-based educational inequality.

Nonetheless, this examination of TFA should reassure proponents, employers, and parents who are in contact with the TFA program. TFA is much more effective in many more subject areas and for many more types of students than the criticism of TFA might suggest. Although this dissertation highlights the fact that TFA's impacts are not uniform, it does show that TFA has positive effects in nearly all subjects and nearly all grades. There are important exceptions that the program must continue to strive to address before TFA can truly claim to be a positive force in *Teaching For All*, however, if TFA is able to perpetuate the successes identified in this dissertation, it should be considered an important and welcome complement to other interventions and training programs serving students in low-income schools.

[38] It is difficult to determine exactly how much students learned in a school year and whether they learned anything in elementary reading with a TFA teacher because the test scores in the administrative data from North Carolina are not vertically integrated. However, using the vertically integrated test scores from the experimental data from Chapter 2 which uses the Iowa Test of Basic Skills (ITBS), it is possible to examine whether the expected amount of learning occurred relative to the national norms for that grade level. The ITBS is scaled so that test scores that remain constant between the beginning and end of a school year represent the typical amount of learning. Any

increase in test scores indicates that a student has learned more than the national norm and has increased their national percentile rankings, and a decrease in test scores indicates that students have learned less than the expected amount and their national percentile ranking has declined. As reported by Decker and colleagues (2004b) in elementary reading, both the TFA and non-TFA students increased their percentile rankings by one point, indicating that in both cases they learned a small amount more than is typical for a given school year, but that there were no differences between the groups in the amount learned on average.

[39] While it is possible that having a TFA teacher for multiple years might have additive effects that would help students to catch up to the state average, a dosage effect of TFA has not yet been tested. However, it is something that I plan to explore in future work.

REFERENCES

Alspaugh, J. W., & Harting, R. D. (1995). Transition effects of school grade-level organization on student achievement. *Journal of Research and Development in Education*, *28*, 145–145.

Amos, J. (2012, August 31). A godsend to low-income communities. *New York Times*. Retrieved from http://www.nytimes.com/roomfordebate/2012/08/30/is-teach-for-america-working/a-godsend-to-low-income-communities

Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., & Walters, C. R. (2010). Inputs and impacts in charter schools: KIPP Lynn. *The American Economic Review*, *100*(2), 239–243.

Antecol, H., Eren, O., & Ozbeklik, S. (Forthcoming). The effect of Teach for America on the distribution of student achievement in primary school: Evidence from a randomized experiment. *Economics of Education Review*.

Antecol, H., Eren, O., & Ozbeklik, S. (2013). *The effect of teach for America on the distribution of student achievement in primary school: Evidence from a randomized experiment*. IZA Discussion Paper No. 6453.

Armecin, G., Behrman, J., Duazo, P., Ghuman, S., Gultiano, S., King, E., & Lee, N. (2006). Early childhood development through an integrated program: evidence from the Philippines. *World Bank Policy Research Working Paper*, (3922).

Arulampalam, W., Naylor, R. A., & Smith, J. (2012). Am I missing something? The effects of absence from class on student performance. *Economics of Education Review*, *31*(4), 363–375.

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*, *122*(3), 1235–1264.

Barahona, G. (2012). *Teach For America announces the schools contributing the most graduates to its 2012 teaching corps* (Press Release). New York, N.Y.: Teach For America. Retrieved from http://www.teachforamerica.org/sites/default/files/20120905_Press.Release_Top.Contributers.pdf

Barnett, W. S. (1996). *Lives in the Balance: Age-27 Benefit-Cost Analysis of the High/Scope Perry Preschool Program. Monographs of the High/Scope Educational Research Foundation, Number Eleven.* ERIC.

Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science*, *333*(6045), 975–978.

Behrman, J. R., Cheng, Y., & Todd, P. E. (2004). Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach. *Review of Economics and Statistics*, *86*(1), 108–132.

Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review*, *28*(5), 629–640.

Belsky, J., Melhuish, E., Barnes, J., Leyland, A. H., Romaniuk, H., & Team, N. E. of S. S. R. (2006). Effects of Sure Start local programmes on children and families: early findings from a quasi-experimental, cross sectional study. *BMJ: British Medical Journal*, 1476–1478.

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, *35*(1), 61–79.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, *1*(4), 289–328.

Bloom, H. S., Raudenbush, S. W., & Weiss, M. (2011). *Estimating variation in program impacts: Theory, practice, and applications*. Working Paper.

Bloom, H. S., Thompson, S., & Unterman, R. (2010). Transforming the high school experience: How New York City's new small schools are boosting student achievement and graduation rates. *MDRC, June*.

Bloom, H. S., & Unterman, R. (2012). *Sustained positive effects on graduation rates produced by New York City's small public high schools of choice*. New York, N.Y.: MDRC.

Boardman, A. E., Davis, O. A., & Sanday, P. R. (1977). A simultaneous equations model of the educational process. *Journal of Public Economics*, *7*(1), 23–49.

Book, C. L., & Freeman, D. J. (1986). Differences in entry characteristics of elementary and secondary teacher candidates. *Journal of Teacher Education*, *37*(2), 47–51.

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research*, *78*(3), 367–409.

Boyd, D. J., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2010). *Recruiting Effective Math Teachers: How Do Math Immersion Teachers Compare?: Evidence from New York City*. National Bureau of Economic Research.

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. (2008). *Who leaves? Teacher attrition and student achievement*. National Bureau of Economic Research.

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.

Boyd, D. J., Lankford, H., Loeb, S., & Wyckoff, J. (2005a). Explaining the short careers of high-achieving teachers in schools with low-performing students. *The American Economic Review*, *95*(2), 166–171.

Boyd, D. J., Lankford, H., Loeb, S., & Wyckoff, J. (2005b). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management*, *24*(1), 113–132.

Bradley, R. H., McKelvey, L. M., & Whiteside-Mansell, L. (2011). Does the quality of stimulation and support in the home environment moderate the effect of early education programs? *Child Development*, *82*(6), 2110–2122.

Bredekamp, S., & Rosegrant, T. (1992). *Reaching Potentials: Appropriate Curriculum and Assessment for Young Children. Volume 1*. ERIC.

Brophy, J., & Evertson, C. M. (1978). Context variables in teaching. *Educational Psychologist*, *12*(3), 310–316.

Brown, B. W., & Saks, D. H. (1986). Measuring the effects of instructional time on student learning: Evidence from the beginning teacher evaluation study. *American Journal of Education*, 480–500.

Bruenig, M. (2012, January 12). *College seniors should reconsider Teach for America*. *The Huffington Post*. Retrieved January 19, 2013, from http://www.huffingtonpost.com/matthew-bruenig/college-seniors-should-re_b_807790.html

Cameron, S. V., & Heckman, J. J. (1993). The nonequivalence of high school equivalents. *Journal of Labor Economics*, *11*(1), 1–47.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families. *Child Development*, *65*(2), 684–698.

Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal*, *14*(3), 368–386.

Carrell, S. E., & Sacerdote, B. (2013). *Late Interventions Matter Too: The Case of College Coaching New Hampshire*. National Bureau of Economic Research.

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

Carter, H., Amrein-Beardsley, A., & Hansen, C. (2011). So NOT amazing! Teach for America corps members' evaluation of the first semester of their teacher preparation program. *The Teachers College Record*, *113*(5), 861–894.

Carter, M. S., & Carter, C. M. (2000). How principals can attract teachers to the middle grades. *Schools in the Middle*, *9*(8), 22–25.

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.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project Star. *The Quarterly Journal of Economics*, *126*(4), 1593.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood*. NBER Working Paper No. 17699.

Chin, A. (2005). Can redistributing teachers across schools raise educational attainment? Evidence from Operation Blackboard in India. *Journal of Development Economics*, *78*(2), 384–405.

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.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007a). *How and why do teacher credentials matter for student achievement?* NBER Working Paper 12828.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007b). 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., & Wheeler, J. (2006). High-poverty schools and the distribution of teachers and principals. *NCL Rev.*, *85*, 1345.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, *23*(7), 13–20.

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.

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, *98*(4), 665.

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.

Croninger, R., & Lee, V. E. (2001). Social capital and dropping out of high school: Benefits to at-risk students of teachers' support and guidance. *The Teachers College Record*, *103*(4), 548–581.

Cumming, G. S., Cumming, D. H., & Redman, C. L. (2006). Scale mismatches in social-ecological systems: causes, consequences, and solutions. *Ecology and Society*, *11*(1), 14.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *The American Economic Review*, *97*(2), 31–47.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, *1*, 697–812.

Currie, J. (2001). Early childhood education programs. *The Journal of Economic Perspectives*, *15*(2), 213–238.

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. (2006). No Child Left Behind and high school reform. *Harvard Educational Review*, *76*(4), 642–667.

Darling-Hammond, L. (2007). Third annual Brown lecture in education research—The flat earth and education: How America's commitment to equity will determine our future. *Educational Researcher*, *36*(6), 318–334.

Darling-Hammond, L. (2011, March 16). Teacher preparation is essential to TFA's future. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2011/03/16/24darling-hammond.h30.html

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.

Darling-Hammond, L., & Sykes, G. (2003). Wanted, A national teacher supply policy for education: The right way to meet the" highly qualified teacher" challenge. *Education Policy Analysis Archives*, *11*, 33.

Decker, P. T., Mayer, D. S., & Glazerman, S. (2004a). *Quality in the classroom: How does Teach For America measure up?* Washington D.C.: Mathematica Policy Research, Inc.

Decker, P. T., Mayer, D. S., & Glazerman, S. (2004b). *The effects of Teach for America on students: Findings from a national evaluation* (No. 8792-750). Princeton, NJ: Mathematica Policy Research, Inc.

Dewey, J. (2004). *Democracy and education*. Courier Dover Publications.

Dobbie, W. (2011). *Teacher characteristics and student achievement: Evidence from Teach For America*. Harvard University.

Dobbie, W., & Fryer Jr, R. G. (2011). *The impact of youth service on future outcomes: Evidence from Teach For America* (No. 17402). National Bureau of Economic Research.

Dobbie, W., & Fryer Jr, R. G. (2013). *The Medium-Term Impacts of High-Achieving Charter Schools on Non-Test Score Outcomes*.

Donaldson, M. L., & 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. *Educational Evaluation and Policy Analysis*, *32*(2), 299–323.

Donaldson, M. L., & Johnson, S. M. (2011). Teach For America teachers: How long do they teach? Why do they leave? *Phi Delta Kappan*, *93*(2), 47–51.

Duflo, E. (2004). The medium run effects of educational expansion: evidence from a large school construction program in Indonesia. *Journal of Development Economics*, *74*(1), 163–197.

Duncan, G. J., & Vandell, D. L. (2011). *Understanding variation in the impacts of human capital Interventions on children and youth*. Working Paper, Irvine Network on Interventions in Development.

Eccles, J. S., Lord, S. E., & Roeser, R. W. (1996). Round holes, square pegs, rocky roads, and sore feet: The impact of stage-environment fit on young adolescents' experiences in schools. *Adolescence: Opportunities and Challenges. Rochester, NY: University of Rochester Press*, *99*, 47–92.

Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. *Research on Motivation in Education*, *3*, 139–186.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: the impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, *48*(2), 90.

Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial and Labor Relations Review*, *48*(3), 547–561.

Eidler, S. (2013, January 3). Deferring six figures on Wall Street for teacher's salary. *New York Times*, p. B3.

Entwisle, D. R., & Alexander, K. L. (1999). Early schooling and social stratification. In *The transition to kindergarten* (pp. 13–38). Baltimore, MD: Brookes Publishing Company.

Erickson, L. D., McDonald, S., & Elder, G. H. (2009). Informal mentors and education: Complementary or compensatory resources? *Sociology of Education*, *82*(4), 344–367.

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*.

ETS. (2014). *Praxis: For Test Takers: About the Tests*. Retrieved August 4, 2014, from https://www.ets.org/praxis/about?WT.ac=praxishome_about_121126

Evans, J. H., & Burck, H. D. (1992). The effects of career education interventions on academic achievement: A meta-analysis. *Journal of Counseling & Development*, *71*(1), 63–68.

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

Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harv. J. on Legis.*, *28*, 465.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. *Holding Schools Accountable: Performance-Based Reform in Education*, 265–298.

Firpo, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, *75*(1), 259–276.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953–973.

Foote, D. (2009). *Relentless pursuit: A year in the trenches with Teach For America*. New York, NY: Vintage Books.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2010). Rethinking response to intervention at middle and high school. *School Psychology Review*, *39*(1), 22–28.

Glazerman, S. (2013). Personal communication.

Glazerman, S., & Grinder, M. (2004). *Mathematica Teach For America Evaluation Public Use Documentation*. Mathematica Policy Research, Inc.

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.

Glazerman, S., Protik, A., Teh, B., Bruch, J., & Max, J. (2013). *Transfer incentives for high-performing teachers: Final results from a multisite randomized experiment*. Mathematica Policy Research.

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

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

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.

Goldstein, D. (2013, September 10). *TFA teachers perform well in a new study -- But teacher experience still matters*. blog. Retrieved December 11, 2013, from http://www.danagoldstein.net/dana_goldstein/2013/09/tfa-teachers-perform-well-in-a-new-study-but-teacher-experience-still-matters.html

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, *66*(3), 361–396.

Gudmundsdottir, S., & Saabar, N. (1991). Cultural Dimensions of the Good Teacher.

Guter, K. (2012, August 31). At least one life changed by Teach for America. *New York Times*. Retrieved from http://www.nytimes.com/roomfordebate/2012/08/30/is-teach-for-america-working/at-least-one-life-changed-by-teach-for-america

Hanushek, E. A. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *The American Economic Review*, 280–288.

Hanushek, E. A. (1972). *Education and Race: An Analysis of the Educational Production Process.* Lexington, MA: D.C. Health.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, *24*(3), 1141–1177.

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

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

Harbison, R. W., & Hanushek, E. A. (1992). *Educational performance of the poor: lessons from rural Northeast Brazil.* Oxford University Press.

Hawkins, J. D., Catalano, R. F., Kosterman, R., Abbott, R., & Hill, K. G. (1999). Preventing adolescent health-risk behaviors by strengthening protection during childhood. *Archives of Pediatrics & Adolescent Medicine*, *153*(3), 226.

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.

Higgins, M., Robison, W., Weiner, J., & Hess, F. (2011). Creating a corps of change agents: What explains the success of Teach for America? *Education Next*, *11*(3), 18–25.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*(2), 371–406.

Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2007). Iowa tests of basic skills (ITBS). *Rolling Meadows, IL: Riverside Publishing*.

Ingersoll, R. M. (1998). The problem of out-of-field teaching. *The Phi Delta Kappan*, *79*(10), 773–776.

Ingersoll, R. M. (1999). The problem of underqualified teachers in American secondary schools. *Educational Researcher*, *28*(2), 26–37.

Ingersoll, R. M. (2004). *Why do high-poverty schools have difficulty staffing their classrooms with qualified teachers?* Center for American Progress, Institute for America's Future.

Ingersoll, R. M., & May, H. (2012). The magnitude, destinations, and determinants of mathematics and science teacher turnover. *Educational Evaluation and Policy Analysis*, *34*(4), 435–464.

Isenberg, E., Max, J., Gleason, P., Potamites, L., Santillano, R., Hock, H., & Hansen, M. (2014). *Access to effective teaching for disadvantaged students: Executive summary.* (No. NCEE 2014-4001). National Center for Education Evaluation and Regional Assistance.

Jackson, E., & Page, M. E. (2013). Estimating the distributional effects of education reforms: A look at Project STAR. *Economics of Education Review*, (32), 92–103.

Jacob, B. A., Lefgren, L., & Sims, D. P. (2010). The persistence of teacher-induced learning. *Journal of Human Resources*, *45*(4), 915–943.

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.

Kemple, J. J., Herlihy, C. M., & Smith, T. J. (2005). Making Progress Toward Graduation: Evidence from the Talent Development High School Model. *MDRC*.

Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, *16*(8), 1412–1425.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences*, *103*(27), 10155–10162.

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.

Kopp, W., & Farr, S. (2011). *A chance to make history: What works and what doesn't in providing an excellent education for all*. New York, NY: PublicAffairs.

Kornrich, S., & Furstenberg, F. (2013). Investing in children: Changes in parental spending on children, 1972–2007. *Demography*, *50*(1), 1–23.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, *114*(2), 497–532.

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/

Lamarche, C. (2007). Voucher program incentives and schooling performance in Colombia: A quantile regression for panel-data approach. *Preprint, University of Oklahoma*.

Lee, V. E., Brooks-Gunn, J., Schnur, E., & Liaw, F.-R. (1990). Are Head Start Effects Sustained? A Longitudinal Follow-up Comparison of Disadvantaged Children Attending Head Start, No Preschool, and Other Preschool Programs. *Child Development*, *61*(2), 495–507.

Lee, V. E., Burkam, D. T., Ready, D. D., Honigman, J., & Meisels, S. J. (2006). Full-Day versus Half-Day Kindergarten: In Which Program Do Children Learn More? *American Journal of Education*, *112*(2), 163–208.

Lee, V. E., & Loeb, S. (1995). Where do Head Start attendees end up? One reason why preschool effects fade out. *Educational Evaluation and Policy Analysis*, *17*(1), 62–82.

Leigh, A. (2010). Estimating teacher effectiveness from two-year changes in students' test scores. *Economics of Education Review*, *29*(3), 480–488. doi:doi: DOI: 10.1016/j.econedurev.2009.10.010

Levin, H. M. (1968). The failure of the public schools and the free market remedy. *The Urban Review*, *2*(7), 32–37.

McAdam, D., & Brandt, C. (2009). Assessing the effects of voluntary youth service: The case of Teach for America. *Social Forces*, *88*(2), 945–969.

McCaffrey, D. F., Hamilton, L. S., Stecher, B. M., Klein, S. P., Bugliari, D., & Robyn, A. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 493–517.

McGuigan, W. M., Katzev, A. R., & Pratt, C. C. (2003). Multi-Level Determinants of Mothers' Engagement in Home Visitation Services*. *Family Relations*, *52*(3), 271–278.

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, *81*(2), 247.

Miller, E. B., Farkas, G., Vandell, D. L., & Duncan, G. J. (2014). Do the Effects of Head Start Vary by Parental Preacademic Stimulation? *Child Development*.

Miller, L. B., & Bizzell, R. P. (1983). Long-term effects of four preschool programs: Sixth, seventh, and eighth grades. *Child Development*, 727–741.

Miner, B. (2010). Looking past the spin: Teach for America. *Rethinking Schools Online*, *24*(3). Retrieved from http://www.rethinkingschools.org/archive/24_03/24_03_TFA.shtml.

Morgan, P. L., Farkas, G., & Hibel, J. (2008). Matthew effects for whom? *Learning Disability Quarterly*, 187–198.

Morris, P., Duncan, G. J., & Clark-Kauffman, E. (2005). Child well-being in an era of welfare reform: The sensitivity of transitions in development to policy change. *Developmental Psychology*, *41*(6), 919.

Mullens, J. E., Murnane, R. J., & Willett, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review, 40*(2), 139–157.

Murnane, R. J. (2013). *US high school graduation rates: Patterns and explanations*. National Bureau of Economic Research.

Murnane, R. J., & Olsen, R. J. (1989). The effect of salaries and opportunity costs on duration in teaching: Evidence from Michigan. *The Review of Economics and Statistics*, 347–352.

Murnane, R. J., & Olsen, R. J. (1990). The effects of salaries and opportunity costs on length of stay in teaching: Evidence from North Carolina. *Journal of Human Resources*, 106–124.

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/

Noell, G. H., & Gansle, K. A. (2009). *Technical report Teach for America teachers' contribution to student achievement in Louisiana in grades 4-9: 2004-2005 to 2006-2007*. Baton Rouge, LA: Louisiana State University.

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

O'Connor, E., & McCartney, K. (2007). Examining teacher–child relationships and achievement as part of an ecological model of development. *American Educational Research Journal, 44*(2), 340–369.

Orfield, G., Losen, D. J., Wald, J., & Swanson, C. B. (2004). *Losing our future: How minority youth are being left behind by the graduation rate crisis*. Cambridge, MA: The Civil Rights Project at Harvard University. Contributors: Advocates for Children of New York, The Civil Society Institute.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. Basic Books.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers.* American Psychological Association.

Pianta, R. C., Steinberg, M. S., & Rollins, K. B. (1995). The first two years of school: Teacher-child relationships and deflections in children's classroom adjustment. *Development and Psychopathology, 7*(02), 295–312.

Pianta, R. C., & Walsh, D. J. (1996). *High-risk children in schools: Constructing sustaining relationships*. Psychology Press.

Raymond, M., Fletcher, S. H., & Luque, J. (2001). *Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Houston, TX.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity: Rising inequality, schools, and children's life chances* (pp. 91–116). New York, N.Y.: Russell Sage Foundation Publications.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, *116*(4), 1092–1153.

Riverside Publishing. (2012). *Iowa Tests of Basic Skills (ITBS_ Forms A, B, and C)*. Retrieved April 2, 2013, from http://www.riversidepublishing.com/products/itbs/details.html

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

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

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4–36.

Roscigno, V. J. (1999). The Black-White Achievement Gap, Family-School Links, and the Importance of Place. *Sociological Inquiry*, *69*(2), 159–186.

Rotherham, A. (2011, February 20). Teach for America: 5 myths that persist 20 years on. *Time Magazine*. Retrieved from http://www.time.com/time/nation/article/0,8599,2047211,00.html

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

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, *125*(1), 175–214.

Rowan, B., Chiang, F.-S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, 256–284.

Rumberger, R. W. (1987). High school dropouts: A review of issues and evidence. *Review of Educational Research*, *57*(2), 101–121.

Rumberger, R. W., & Lamb, S. P. (2003). The early employment and further education experiences of high school dropouts: A comparative study of the United States and Australia. *Economics of Education Review*, *22*(4), 353–366.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry*, *57*(3), 316.

Sameroff, A. J., & Chandler, M. J. (1975). Reproductive risk and the continuum of caretaking casualty. *Review of Child Development Research*, *4*, 187–244.

Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*(1), 57–67.

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

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*. Center for Research and Evaluation Office of Accountability: Charlotte-Mecklenburg Schools.

Schorr, J. (1993). Class action: What Clinton's National Service Program could learn from "Teach for America." *Phi Delta Kappan*, *75*(4), 315–318.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–23.

Simon, S. (2012, August 16). Has Teach for America betrayed its mission? *Reuters*. Retrieved from http://www.reuters.com/article/2012/08/16/us-usa-education-teachforamerica-idUSBRE87F05O20120816

Skandera, H. (2012). *K-3 Plus Education annual report for school year 2011-2012*. New Mexico Public Education Department.

Socol, I. D. (2008, December 11). Ira David Socol on Teach for America, KIPP Schools, and Reforming Education. Retrieved from http://www.openeducation.net/2008/12/11/ira-david-socol-on-teach-for-america-kipp-schools-and-reforming-education/

Spencer, J. T. (2012). Abandoning Superman. *Phi Delta Kappan*, *94*(1), 72–73.

Stanton-Salazar, R. D., & Spina, S. U. (2003). Informal mentors and role models in the lives of urban Mexican-origin adolescents. *Anthropology & Education Quarterly*, *34*(3), 231–254.

Strategic Data Project. (2012). *SDP Human Capital Diagnostic: Los Angeles Unified*. Center for Education Policy Research: Harvard University.

Strauss, R. P., & Sawyer, E. A. (1986). Some new evidence on teacher and student competencies. *Economics of Education Review*, *5*(1), 41–48.

Suh, S., Suh, J., & Houston, I. (2007). Predictors of Categorical At-Risk High School Dropouts. *Journal of Counseling & Development*, *85*(2), 196–203.

Teach For America. (2012). *Investing In Innovation (i3) Fund: Scaling Teach For America: Growing the talent force working to ensure all our nation's students have access to a quality education*. Washington D.C.: U.S. Department of Education Office of Innovation and Improvement. Retrieved from http://www2.ed.gov/programs/innovation/2010/narratives/u396a100015.pdf

Teach For America. (2013a). *Coaching*. *Teach For America*. Retrieved January 22, 2013, from http://www.teachforamerica.org/why-teach-for-america/training-and-support/coaching

Teach For America. (2013b). *School leadership initiative*. Retrieved January 28, 2013, from http://www.teachforamerica.org/corps-member-and-alumni-resources/alumni-leadership-initiatives/educational-leadership/school-leadership-initiative

Teach For America. (2013c, May 31). Teach For America: Annual Letter 2012-2013. Retrieved from http://www.teachforamerica.org/sites/default/files/teach_for_america_fy12-13_annual_letter.pdf

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

teachforamerica.org. (2013b). *Where We Work*. *Teach For America*. 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. (2014). *New Sites*. *Teach For America*. Retrieved July 7, 2014, from http://www.teachforamerica.org/where-we-work/new-sites

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

Turner, H. M., Goodman, D., Adachi, E., Brite, J., & Decker, L. (2012). *Evaluation of Teach For America in Texas schools*. Edvance Research, Inc.

United States Census Bureau. (2011, June 27). *Facts for features: Back to school: 2011-2012*. *census.gov*. Retrieved February 1, 2013, from http://www.census.gov/newsroom/releases/archives/facts_for_features_special_editions/cb11-ff15.html

Useem, E., & Neild, R. C. (2001). Teacher Staffing in the School District of Philadelphia: A Report to the Community.

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.

Vasquez Heilig, J., Cole, H. A., & Springel, M. A. (2010). Alternative certification and the Teach for America: the search for high quality teachers. *Kansas Journal of Law & Public Policy*, *20*, 388.

Vasquez Heilig, J., & Jez, S. J. (2010). *Teach For America: A review of the evidence* (Policy Brief). Great Lakes Center for Education Research and Practice.

Veltri, B. T. (2010). *Learning on other people's kids: Becoming a Teach For America teacher*. Charlotte, NC: Information Age Publishing Incorporated.

Walker, H. M., & Shinn, M. R. (2002). Structuring school-based interventions to achieve integrated primary, secondary, and tertiary prevention goals for safe and effective schools. *Interventions for Academic and Behavior Problems II: Preventive and Remedial Approaches*, 1–25.

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.

Watamura, S. E., Phillips, D. A., Morrissey, T. W., McCartney, K., & Bub, K. (2011). Double Jeopardy: Poorer Social-Emotional Outcomes for Children in the NICHD SECCYD Experiencing Home and Child-Care Environments That Confer Risk. *Child Development*, *82*(1), 48–65.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89–122.

Weiss, M. J., Bloom, H. S., & Brock, T. (2013). *A conceptual framework for studying the sources of variation in program effects* (MDRC Working Paper on Research Methodology).

Wilson, S. M., Shulman, L. S., & Richert, A. E. (1987). "150 different ways" of knowing: Representations of knowledge in teaching. In *Exploring teachers' thinking* (pp. 104–124). Sussex, England: Holt, Rinehart, & Winston.

Woodward, J., & Baxter, J. (1997). The effects of an innovative approach to mathematics on academically low-achieving students in mainstreamed settings. *Exceptional Children*, *63*, 373–388.

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.

APPENDICES

Appendix Table 3.1

Impact of TFA on same-year math and reading tests, grades 4-8, with school-grade-year fixed effects

	Math					Reading				
	Grade 4	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**	**Grade 4**	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**
TFA	**0.052*****	**0.062*****	**0.050*****	**0.041*****	**0.048*****	-0.015	**-0.034***	0.004	0.011	**0.051*****
	(0.015)	(0.015)	(0.010)	(0.009)	(0.010)	(0.016)	(0.017)	(0.010)	(0.011)	(0.010)
Lagged Reading Score	0.081***	0.071***	0.087***	0.089***	0.098***	0.510***	0.511***	0.512***	0.545***	0.532***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Reading Score is Missing	-0.034***	-0.005	-0.018**	0.026***	0.010	-0.702***	-0.739***	-0.828***	-0.762***	-0.748***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)
Lagged Math Score	0.554***	0.576***	0.554***	0.591***	0.585***	0.130***	0.134***	0.128***	0.130***	0.141***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Math Score is Missing	-0.177***	-0.267***	-0.272***	-0.356***	-0.372***	0.586***	0.546***	0.592***	0.482***	0.449***
	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)
Black	-0.202***	-0.176***	-0.185***	-0.147***	-0.157***	-0.193***	-0.187***	-0.196***	-0.171***	-0.214***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Hispanic	-0.027***	-0.039***	-0.038***	-0.035***	-0.039***	-0.061***	-0.064***	-0.058***	-0.041***	-0.080***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Asian	0.119***	0.128***	0.151***	0.154***	0.172***	-0.018***	-0.042***	-0.005	-0.008*	-0.041***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
American Indian	-0.105***	-0.082***	-0.102***	-0.070***	-0.093***	-0.107***	-0.108***	-0.105***	-0.087***	-0.115***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Other Race	-0.064***	-0.048***	-0.055***	-0.041***	-0.046***	-0.050***	-0.038***	-0.039***	-0.015***	-0.037***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Female	-0.042***	-0.046***	-0.030***	-0.024***	-0.029***	0.047***	0.032***	0.060***	0.038***	0.026***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.068***	0.068***	0.074***	0.063***	0.074***	0.083***	0.095***	0.094***	0.099***	0.107***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed: College or more	0.165***	0.159***	0.164***	0.137***	0.140***	0.180***	0.179***	0.182***	0.164***	0.164***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed missing	0.090***	0.130***	0.163***	0.131***	0.133***	0.084***	0.150***	0.171***	0.156***	0.157***
	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)	(0.006)	(0.004)	(0.004)	(0.003)	(0.003)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Education is HS or less; Standard errors in parentheses.

Appendix Table 3.1 continued

Impact of TFA on same-year math and reading tests, grades 4-8, with school-grade-year fixed effects

	Math					Reading				
	Grade 4	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**	**Grade 4**	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**
Student has disability	-0.212***	-0.206***	-0.230***	-0.194***	-0.199***	-0.261***	-0.262***	-0.283***	-0.266***	-0.263***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Student is gifted	0.486***	0.451***	0.432***	0.421***	0.345***	0.417***	0.355***	0.354***	0.284***	0.272***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Gifted status missing	-0.506***	-0.713***	-0.587***	-0.512***	-0.493***	-0.259**	-0.446***	-0.508***	-0.666***	-0.597***
	(0.057)	(0.054)	(0.054)	(0.052)	(0.049)	(0.087)	(0.094)	(0.092)	(0.087)	(0.086)
Student ever IDed as LEP	-0.048***	-0.037***	-0.055***	-0.022***	-0.018***	-0.139***	-0.164***	-0.140***	-0.133***	-0.186***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LEP status missing	-0.111*	-0.120**	-0.250***	-0.221***	-0.468***	-0.348***	-0.283***	-0.359***	-0.266***	-0.447***
	(0.050)	(0.044)	(0.045)	(0.043)	(0.041)	(0.057)	(0.050)	(0.051)	(0.050)	(0.046)
Student over age for grade	-0.127***	-0.129***	-0.128***	-0.115***	-0.143***	-0.122***	-0.115***	-0.116***	-0.109***	-0.136***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.050***	0.057***	0.058***	0.070***	0.065***	0.051***	0.046***	0.030***	0.037***	0.040***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Student has repeated grade	-0.579***	-0.529***	-0.448***	-0.374***	-0.447***	-0.665***	-0.676***	-0.490***	-0.449***	-0.460***
	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)
Repeater status missing	-0.275	-0.483**	-1.020***	-0.435***	-0.275**	-0.686***	-1.161***	-1.012***	-1.048***	-0.199*
	(0.160)	(0.184)	(0.152)	(0.110)	(0.089)	(0.192)	(0.201)	(0.158)	(0.140)	(0.101)
Constant	-0.035***	-0.033***	-0.022***	-0.032***	-0.009***	-0.068***	-0.054***	-0.048***	-0.032***	-0.008***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.619	0.642	0.652	0.667	0.654	0.588	0.588	0.605	0.604	0.606
Observations	1259636	1253760	1253676	1249201	1231416	1251642	1246125	1248636	1245332	1228326

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Education is HS or less; Standard errors in parentheses.

Appendix Table 3.2

Impact of TFA on same-year tests, high school, with school-subject-year fixed effects

	Eng 1	Alg 1	Alg 2	Geom	Chem	Bio	Phy	Psci	Elp	Ush	Civ
TFA	**0.036*****	**0.103*****	**0.076*****	**0.163*****	**0.355*****	**0.172*****	0.138	**0.245*****	**-0.971*****	**0.090****	**0.051*****
	(0.010)	(0.011)	(0.021)	(0.018)	(0.044)	(0.012)	(0.165)	(0.025)	(0.123)	(0.028)	(0.013)
8th grade reading score	0.440***	0.092***	0.037***	0.067***	0.140***	0.318***	0.142***	0.206***	0.390***	0.380***	0.410***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
8th grade reading score is missing	-0.573***	0.054***	-0.141***	-0.182***	0.154*	-0.205***	0.526**	-0.016	-0.257***	-0.177***	-0.247***
	(0.009)	(0.010)	(0.021)	(0.019)	(0.069)	(0.012)	(0.165)	(0.019)	(0.033)	(0.018)	(0.013)
8th grade math score	0.168***	0.532***	0.601***	0.621***	0.498***	0.290***	0.525***	0.430***	0.161***	0.144***	0.253***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
8th grade math score is missing	0.361***	-0.436***	0.274***	0.218***	0.293***	0.174***	0.315	-0.147***	0.431***	0.317***	0.336***
	(0.009)	(0.010)	(0.021)	(0.019)	(0.069)	(0.012)	(0.166)	(0.019)	(0.033)	(0.018)	(0.013)
Black	-0.198***	-0.127***	-0.158***	-0.262***	-0.209***	-0.296***	-0.380***	-0.261***	-0.355***	-0.303***	-0.190***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.009)	(0.003)	(0.003)	(0.002)	(0.002)
Hispanic	-0.085***	-0.019***	-0.024***	-0.066***	-0.028***	-0.082***	-0.133***	-0.066***	-0.143***	-0.011*	-0.077***
	(0.003)	(0.004)	(0.005)	(0.004)	(0.008)	(0.004)	(0.018)	(0.006)	(0.007)	(0.005)	(0.005)
Asian	-0.025***	0.151***	0.272***	0.155***	0.186***	0.061***	0.053***	0.063***	-0.068***	0.052***	0.012*
	(0.004)	(0.005)	(0.005)	(0.005)	(0.007)	(0.005)	(0.012)	(0.009)	(0.008)	(0.006)	(0.006)
American Indian	-0.142***	-0.089***	-0.110***	-0.127***	-0.118***	-0.167***	-0.243***	-0.168***	-0.207***	-0.175***	-0.151***
	(0.006)	(0.006)	(0.009)	(0.008)	(0.015)	(0.007)	(0.039)	(0.010)	(0.012)	(0.009)	(0.009)
Other Race	-0.028***	-0.022***	-0.054***	-0.082***	-0.052***	-0.054***	0.008	-0.068***	-0.070***	-0.059***	-0.059***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.010)	(0.005)	(0.023)	(0.007)	(0.011)	(0.006)	(0.005)
Female	0.137***	-0.005***	0.024***	-0.077***	-0.114***	-0.123***	-0.299***	-0.168***	-0.170***	-0.238***	-0.174***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed: Some college/trade	0.080***	0.035***	0.028***	0.054***	0.075***	0.084***	0.080***	0.070***	0.189***	0.122***	0.034***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.002)	(0.011)	(0.003)	(0.003)	(0.003)	(0.003)
Parent Ed: College or more	0.139***	0.057***	0.095***	0.111***	0.144***	0.129***	0.195***	0.091***	0.237***	0.194***	0.071***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)	(0.011)	(0.003)	(0.003)	(0.003)	(0.003)
Parent Ed is missing	0.173***	0.087***	0.213***	0.227***	0.343***	0.232***	0.194***	0.098***	-0.244***	0.208***	0.143***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.012)	(0.004)	(0.030)	(0.006)	(0.038)	(0.005)	(0.004)

* p<0.05 ** p<0.01 *** p<0.001; Standard errors in parentheses.

Appendix Table 3.2 continued

Impact of TFA on same-year tests, high school, with school-subject-year fixed effects

	Eng 1	**Alg 1**	**Alg 2**	**Geom**	**Chem**	**Bio**	**Phy**	**Psci**	**Elp**	**Ush**	**Civ**
Student has disability	-0.327***	-0.165***	-0.062***	-0.086***	-0.054***	-0.160***	-0.039**	-0.176***	-0.348***	-0.149***	-0.086***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.006)	(0.002)	(0.013)	(0.003)	(0.004)	(0.003)	(0.002)
Disability is missing	-0.417***	-0.021	0.005	-0.210**	0.001	-0.265***	0.211	-0.094	0.300	-0.169***	-0.358***
	(0.032)	(0.038)	(0.057)	(0.068)	(0.081)	(0.045)	(0.127)	(0.067)	(0.510)	(0.039)	(0.042)
Student is gifted	0.326***	0.257***	0.365***	0.341***	0.399***	0.337***	0.381***	0.348***	0.474***	0.412***	0.241***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.006)	(0.003)	(0.004)	(0.002)	(0.003)
Gifted status missing	-0.242**	-0.241***	-0.199	0.000	0.000	0.037	0.000	-0.272	0.000	-0.137	-0.396***
	(0.083)	(0.070)	(0.115)	(.)	(.)	(0.120)	(.)	(0.185)	(.)	(0.112)	(0.090)
Student ever IDed as LEP	-0.223***	-0.023***	0.054***	0.019***	0.025**	-0.146***	0.025	-0.101***	-0.315***	-0.100***	-0.185***
	(0.003)	(0.004)	(0.005)	(0.004)	(0.008)	(0.004)	(0.018)	(0.006)	(0.008)	(0.005)	(0.005)
LEP status missing	-0.505***	0.023	0.294**	0.175	0.306	-0.380***	0.813*	-0.113	0.000	-0.270**	-0.341***
	(0.060)	(0.060)	(0.094)	(0.095)	(0.334)	(0.092)	(0.343)	(0.110)	(.)	(0.083)	(0.066)
Student over age for grade	-0.202***	-0.200***	-0.159***	-0.163***	-0.147***	-0.240***	-0.105***	-0.240***	-0.265***	-0.254***	-0.218***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)	(0.002)	(0.009)	(0.002)	(0.003)	(0.002)	(0.002)
Student under age for grade	0.109***	0.113***	0.168***	0.144***	0.221***	0.166***	0.229***	0.161***	0.161***	0.170***	0.157***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.009)	(0.006)	(0.017)	(0.010)	(0.010)	(0.007)	(0.007)
Student has repeated grade	-0.421***	-0.494***	-0.613***	-0.527***	-0.578***	-0.621***	0.295**	-0.650***	-0.729***	-0.686***	-0.526***
	(0.002)	(0.003)	(0.009)	(0.007)	(0.026)	(0.005)	(0.109)	(0.009)	(0.010)	(0.009)	(0.006)
Repeater status missing	-0.249***	-0.371***	-0.427***	-0.649***	-0.242	-0.632***	-0.019	-0.732***	-0.812***	-0.597***	-0.361***
	(0.022)	(0.024)	(0.044)	(0.086)	(0.377)	(0.037)	(0.355)	(0.056)	(0.143)	(0.045)	(0.036)
Constant	-0.004*	0.018***	-0.384***	-0.253***	-0.580***	0.015***	-0.924***	0.252***	-0.070***	-0.053***	0.065***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.002)	(0.013)	(0.003)	(0.004)	(0.003)	(0.003)
R-squared	0.594	0.434	0.397	0.491	0.317	0.488	0.275	0.398	0.443	0.370	0.517
Observations	1270188	1000489	821271	800758	400506	1126915	92116	555098	380616	854769	629854

* p<0.05 ** p<0.01 *** p<0.001; Standard errors in parentheses.

Appendix Table 3.3

Impact of TFA on same-year tests based on baseline proficiency and school level, with school-subject-year fixed effects

	Math	ELA
TFA	**0.052****	-0.030
	(0.020)	(0.020)
TFA*Middle School	-0.006	**0.052***
	(0.023)	(0.023)
TFA*High School	0.043	**0.058***
	(0.023)	(0.026)
Bottom Tercile baseline test (math or reading)	-0.582***	-0.597***
	(0.001)	(0.001)
Top Tercile baseline test (math or reading)	0.595***	0.498***
	(0.001)	(0.002)
TFA*Bottom tercile baseline test	0.012	0.020
	(0.022)	(0.023)
TFA*Top tercile baseline test	0.040	**0.113*****
	(0.030)	(0.031)
TFA*missing baseline test	**-0.062***	0.022
	(0.027)	(0.028)
Middle School*Bottom tercile baseline test	0.071***	0.025***
	(0.002)	(0.002)
Middle School*Top tercile baseline test	0.039***	-0.011***
	(0.002)	(0.002)
High School*Bottom tercile baseline test	0.184***	0.124***
	(0.002)	(0.002)
High School*Top tercile baseline test	-0.002	**0.018*****
	(0.002)	(0.003)
TFA*Middle School*Bottom Tercile baseline test	-0.003	-0.023
	(0.026)	(0.026)
TFA*Middle School*Top Tercile baseline test	0.018	-0.016
	(0.035)	(0.035)
TFA*High School*Bottom Tercile baseline test	0.028	0.026
	(0.027)	(0.031)
TFA*High School*Top Tercile baseline test	-0.038	**-0.115****
	(0.038)	(0.042)
Black	-0.248***	-0.268***
	(0.001)	(0.001)
Hispanic	-0.062***	-0.097***
	(0.002)	(0.002)
Asian	0.205***	0.001
	(0.002)	(0.002)
American Indian	-0.134***	-0.155***
	(0.003)	(0.003)
Other Race	-0.092***	-0.080***
	(0.002)	(0.002)

* p<0.05 ** p<0.01 *** p<0.001; Baseline tests are lagged test scores in grades 4-8 and 8th grade tests in grades 9-12; Omitted categories are: White; Parent Ed. High Sch. or less; Less than median baseline test score, Elementary School; Standard errors in parentheses.

Appendix Table 3.3 continued

Impact of TFA on same-year tests based on baseline proficiency and school level, with school-subject-year fixed effects

	Math	ELA
Female	-0.038***	0.052***
Parent Ed: Some college/trade school	0.048***	0.081***
Parent Ed: College or more	0.106***	0.153***
Parent Ed: Missing	0.048***	0.037***
Student has disability	-0.270***	-0.376***
Disability status is missing	-0.034	-0.431***
Student is gifted	0.584***	0.560***
Gifted status missing	-0.596***	-0.491***
Student ever IDed as LEP	-0.098***	-0.256***
LEP status missing	-0.301***	-0.421***
Student over age for grade	-0.219***	-0.221***
Student under age for grade	0.091***	0.048***
Student has repeated grade	-0.687***	-0.639***
Repeater status missing	-0.653***	-0.813***
	(0.032)	(0.045)
Constant	-0.026***	0.069***
	(0.001)	(0.001)
R-squared	0.497	0.501
Observations	5261101	4533172

* p<0.05 ** p<0.01 *** p<0.001; Baseline tests are lagged test scores in grades 4-8 and 8th grade tests in grades 9-12; Omitted categories are: White; Parent Ed. High Sch. or less; Less than median baseline test score, Elementary School; Standard errors in parentheses.

Appendix Table 4.1

Impact of TFA on same-year math and reading tests, grades 4-8, with school-grade-year fixed effects

	Math					Reading				
	Grade 4	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**	**Grade 4**	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**
TFA	**0.052*****	**0.062*****	**0.050*****	**0.041*****	**0.048*****	-0.015	**-0.034***	0.004	0.011	**0.051*****
	(0.015)	(0.015)	(0.010)	(0.009)	(0.010)	(0.016)	(0.017)	(0.010)	(0.011)	(0.010)
Lagged Reading Score	0.081***	0.071***	0.087***	0.089***	0.098***	0.510***	0.511***	0.512***	0.545***	0.532***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Reading Score is Missing	-0.034***	-0.005	-0.018**	0.026***	0.010	-0.702***	-0.739***	-0.828***	-0.762***	-0.748***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)
Lagged Math Score	0.554***	0.576***	0.554***	0.591***	0.585***	0.130***	0.134***	0.128***	0.130***	0.141***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Math Score is Missing	-0.177***	-0.267***	-0.272***	-0.356***	-0.372***	0.586***	0.546***	0.592***	0.482***	0.449***
	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)
Black	-0.202***	-0.176***	-0.185***	-0.147***	-0.157***	-0.193***	-0.187***	-0.196***	-0.171***	-0.214***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Hispanic	-0.027***	-0.039***	-0.038***	-0.035***	-0.039***	-0.061***	-0.064***	-0.058***	-0.041***	-0.080***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Asian	0.119***	0.128***	0.151***	0.154***	0.172***	-0.018***	-0.042***	-0.005	-0.008*	-0.041***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
American Indian	-0.105***	-0.082***	-0.102***	-0.070***	-0.093***	-0.107***	-0.108***	-0.105***	-0.087***	-0.115***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Other Race	-0.064***	-0.048***	-0.055***	-0.041***	-0.046***	-0.050***	-0.038***	-0.039***	-0.015***	-0.037***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Female	-0.042***	-0.046***	-0.030***	-0.024***	-0.029***	0.047***	0.032***	0.060***	0.038***	0.026***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.068***	0.068***	0.074***	0.063***	0.074***	0.083***	0.095***	0.094***	0.099***	0.107***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed: College or more	0.165***	0.159***	0.164***	0.137***	0.140***	0.180***	0.179***	0.182***	0.164***	0.164***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed missing	0.090***	0.130***	0.163***	0.131***	0.133***	0.084***	0.150***	0.171***	0.156***	0.157***
	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)	(0.006)	(0.004)	(0.004)	(0.003)	(0.003)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Education is HS or less; Standard errors in parentheses.

Appendix Table 4.1 continued

Impact of TFA on same-year math and reading tests, grades 4-8, with school-grade-year fixed effects

	Math					Reading				
	Grade 4	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**	**Grade 4**	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**
Student has disability	-0.212***	-0.206***	-0.230***	-0.194***	-0.199***	-0.261***	-0.262***	-0.283***	-0.266***	-0.263***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Student is gifted	0.486***	0.451***	0.432***	0.421***	0.345***	0.417***	0.355***	0.354***	0.284***	0.272***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Gifted status missing	-0.506***	-0.713***	-0.587***	-0.512***	-0.493***	-0.259**	-0.446***	-0.508***	-0.666***	-0.597***
	(0.057)	(0.054)	(0.054)	(0.052)	(0.049)	(0.087)	(0.094)	(0.092)	(0.087)	(0.086)
Student ever IDed as LEP	-0.048***	-0.037***	-0.055***	-0.022***	-0.018***	-0.139***	-0.164***	-0.140***	-0.133***	-0.186***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LEP status missing	-0.111*	-0.120**	-0.250***	-0.221***	-0.468***	-0.348***	-0.283***	-0.359***	-0.266***	-0.447***
	(0.050)	(0.044)	(0.045)	(0.043)	(0.041)	(0.057)	(0.050)	(0.051)	(0.050)	(0.046)
Student over age for grade	-0.127***	-0.129***	-0.128***	-0.115***	-0.143***	-0.122***	-0.115***	-0.116***	-0.109***	-0.136***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.050***	0.057***	0.058***	0.070***	0.065***	0.051***	0.046***	0.030***	0.037***	0.040***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Student has repeated grade	-0.579***	-0.529***	-0.448***	-0.374***	-0.447***	-0.665***	-0.676***	-0.490***	-0.449***	-0.460***
	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)
Repeater status missing	-0.275	-0.483**	-1.020***	-0.435***	-0.275**	-0.686***	-1.161***	-1.012***	-1.048***	-0.199*
	(0.160)	(0.184)	(0.152)	(0.110)	(0.089)	(0.192)	(0.201)	(0.158)	(0.140)	(0.101)
Constant	-0.035***	-0.033***	-0.022***	-0.032***	-0.009***	-0.068***	-0.054***	-0.048***	-0.032***	-0.008***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.619	0.642	0.652	0.667	0.654	0.588	0.588	0.605	0.604	0.606
Observations	1259636	1253760	1253676	1249201	1231416	1251642	1246125	1248636	1245332	1228326

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Education is HS or less; Standard errors in parentheses.

Appendix Table 4.2

Impact of TFA on same-year tests, high school, with school-subject-year fixed effects

	Eng 1	Alg 1	Alg 2	Geom	Chem	Bio	Phy	Psci	Elp	Ush	Civ
TFA	**0.036*****	**0.103*****	**0.076*****	**0.163*****	**0.355*****	**0.172*****	0.138	**0.245*****	**-0.971*****	**0.090****	**0.051*****
	(0.010)	(0.011)	(0.021)	(0.018)	(0.044)	(0.012)	(0.165)	(0.025)	(0.123)	(0.028)	(0.013)
8th grade reading score	0.440***	0.092***	0.037***	0.067***	0.140***	0.318***	0.142***	0.206***	0.390***	0.380***	0.410***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
8th grade reading score is missing	-0.573***	0.054***	-0.141***	-0.182***	0.154*	-0.205***	0.526**	-0.016	-0.257***	-0.177***	-0.247***
	(0.009)	(0.010)	(0.021)	(0.019)	(0.069)	(0.012)	(0.165)	(0.019)	(0.033)	(0.018)	(0.013)
8th grade math score	0.168***	0.532***	0.601***	0.621***	0.498***	0.290***	0.525***	0.430***	0.161***	0.144***	0.253***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
8th grade math score is missing	0.361***	-0.436***	0.274***	0.218***	0.293***	0.174***	0.315	-0.147***	0.431***	0.317***	0.336***
	(0.009)	(0.010)	(0.021)	(0.019)	(0.069)	(0.012)	(0.166)	(0.019)	(0.033)	(0.018)	(0.013)
Black	-0.198***	-0.127***	-0.158***	-0.262***	-0.209***	-0.296***	-0.380***	-0.261***	-0.355***	-0.303***	-0.190***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.009)	(0.003)	(0.003)	(0.002)	(0.002)
Hispanic	-0.085***	-0.019***	-0.024***	-0.066***	-0.028***	-0.082***	-0.133***	-0.066***	-0.143***	-0.011*	-0.077***
	(0.003)	(0.004)	(0.005)	(0.004)	(0.008)	(0.004)	(0.018)	(0.006)	(0.007)	(0.005)	(0.005)
Asian	-0.025***	0.151***	0.272***	0.155***	0.186***	0.061***	0.053***	0.063***	-0.068***	0.052***	0.012*
	(0.004)	(0.005)	(0.005)	(0.005)	(0.007)	(0.005)	(0.012)	(0.009)	(0.008)	(0.006)	(0.006)
American Indian	-0.142***	-0.089***	-0.110***	-0.127***	-0.118***	-0.167***	-0.243***	-0.168***	-0.207***	-0.175***	-0.151***
	(0.006)	(0.006)	(0.009)	(0.008)	(0.015)	(0.007)	(0.039)	(0.010)	(0.012)	(0.009)	(0.009)
Other Race	-0.028***	-0.022***	-0.054***	-0.082***	-0.052***	-0.054***	0.008	-0.068***	-0.070***	-0.059***	-0.059***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.010)	(0.005)	(0.023)	(0.007)	(0.011)	(0.006)	(0.005)
Female	0.137***	-0.005***	0.024***	-0.077***	-0.114***	-0.123***	-0.299***	-0.168***	-0.170***	-0.238***	-0.174***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.080***	0.035***	0.028***	0.054***	0.075***	0.084***	0.080***	0.070***	0.189***	0.122***	0.034***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.002)	(0.011)	(0.003)	(0.003)	(0.003)	(0.003)
Parent Ed: College or more	0.139***	0.057***	0.095***	0.111***	0.144***	0.129***	0.195***	0.091***	0.237***	0.194***	0.071***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)	(0.011)	(0.003)	(0.003)	(0.003)	(0.003)

* p<0.05 ** p<0.01 *** p<0.001; Standard errors in parentheses.

Appendix Table 4.2 continued

Impact of TFA on same-year tests, high school, with school-subject-year fixed effects

	Eng 1	Alg 1	Alg 2	Geom	Chem	Bio	Phy	Psci	Elp	Ush	Civ
Parent Ed is missing	0.173***	0.087***	0.213***	0.227***	0.343***	0.232***	0.194***	0.098***	-0.244***	0.208***	0.143***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.012)	(0.004)	(0.030)	(0.006)	(0.038)	(0.005)	(0.004)
Student has disability	-0.327***	-0.165***	-0.062***	-0.086***	-0.054***	-0.160***	-0.039**	-0.176***	-0.348***	-0.149***	-0.086***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.006)	(0.002)	(0.013)	(0.003)	(0.004)	(0.003)	(0.002)
Disability is missing	-0.417***	-0.021	0.005	-0.210**	0.001	-0.265***	0.211	-0.094	0.300	-0.169***	-0.358***
	(0.032)	(0.038)	(0.057)	(0.068)	(0.081)	(0.045)	(0.127)	(0.067)	(0.510)	(0.039)	(0.042)
Student is gifted	0.326***	0.257***	0.365***	0.341***	0.399***	0.337***	0.381***	0.348***	0.474***	0.412***	0.241***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.006)	(0.003)	(0.004)	(0.002)	(0.003)
Gifted status missing	-0.242**	-0.241***	-0.199	0.000	0.000	0.037	0.000	-0.272	0.000	-0.137	-0.396***
	(0.083)	(0.070)	(0.115)	(.)	(.)	(0.120)	(.)	(0.185)	(.)	(0.112)	(0.090)
Student ever IDed as LEP	-0.223***	-0.023***	0.054***	0.019***	0.025**	-0.146***	0.025	-0.101***	-0.315***	-0.100***	-0.185***
	(0.003)	(0.004)	(0.005)	(0.004)	(0.008)	(0.004)	(0.018)	(0.006)	(0.008)	(0.005)	(0.005)
LEP status missing	-0.505***	0.023	0.294**	0.175	0.306	-0.380***	0.813*	-0.113	0.000	-0.270**	-0.341***
	(0.060)	(0.060)	(0.094)	(0.095)	(0.334)	(0.092)	(0.343)	(0.110)	(.)	(0.083)	(0.066)
Student over age for grade	-0.202***	-0.200***	-0.159***	-0.163***	-0.147***	-0.240***	-0.105***	-0.240***	-0.265***	-0.254***	-0.218***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)	(0.002)	(0.009)	(0.002)	(0.003)	(0.002)	(0.002)
Student under age for grade	0.109***	0.113***	0.168***	0.144***	0.221***	0.166***	0.229***	0.161***	0.161***	0.170***	0.157***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.009)	(0.006)	(0.017)	(0.010)	(0.010)	(0.007)	(0.007)
Student has repeated grade	-0.421***	-0.494***	-0.613***	-0.527***	-0.578***	-0.621***	0.295**	-0.650***	-0.729***	-0.686***	-0.526***
	(0.002)	(0.003)	(0.009)	(0.007)	(0.026)	(0.005)	(0.109)	(0.009)	(0.010)	(0.009)	(0.006)
Repeater status missing	-0.249***	-0.371***	-0.427***	-0.649***	-0.242	-0.632***	-0.019	-0.732***	-0.812***	-0.597***	-0.361***
	(0.022)	(0.024)	(0.044)	(0.086)	(0.377)	(0.037)	(0.355)	(0.056)	(0.143)	(0.045)	(0.036)
Constant	-0.004*	0.018***	-0.384***	-0.253***	-0.580***	0.015***	-0.924***	0.252***	-0.070***	-0.053***	0.065***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.002)	(0.013)	(0.003)	(0.004)	(0.003)	(0.003)
R-squared	0.594	0.434	0.397	0.491	0.317	0.488	0.275	0.398	0.443	0.370	0.517
Observations	1270188	1000489	821271	800758	400506	1126915	92116	555098	380616	854769	629854

* p<0.05 ** p<0.01 *** p<0.001; Standard errors in parentheses.

Appendix Table 4.3

Impact of TFA on achievement, elementary, middle, and high school, controlling for teacher Praxis scores

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.029****	**-0.023***	**0.041*****	**0.022*****	**0.026****	**0.093*****	**0.179*****	0.024
	(0.010)	(0.011)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Mean Praxis Score	**0.016*****	**0.004*****	**0.009*****	**0.003*****	**0.014*****	**0.013*****	**0.027*****	**0.033*****
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Praxis Score Missing	-0.027***	-0.014***	-0.016***	-0.011***	-0.025***	-0.065***	-0.046***	-0.043***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.001	-0.088***	-0.211***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.565***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score Missing	-0.343***	0.478***	-0.327***	0.514***	0.361***	-0.106***	0.138***	0.340***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.197***	-0.186***	-0.283***	-0.279***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.164***	-0.176***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.124***	0.166***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.155***	0.139***	0.159***	0.173***	0.160***	0.163***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.320***	-0.397***	-0.208***	-0.271***	-0.326***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.411***	-0.075*	-0.055	-0.224***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.615***	0.400***	0.304***	0.325***	0.335***	0.356***	0.369***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.617***	-0.429***	-0.540***	-0.591***	-0.245**	-0.360***	-0.201*	-0.288***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.144***	-0.245***	-0.318***	-0.360***	-0.500***	0.085	-0.215**	-0.300***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.169***	-0.168***	-0.129***	-0.121***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.630***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.279*	-0.824***	-0.461***	-0.604***	-0.249***	-0.407***	-0.674***	-0.476***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.080***	0.016***	-0.016***	-0.027***	0.002	-0.143***	-0.025***	-0.011***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.534	0.501	0.658	0.605	0.594	0.429	0.416	0.430
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.4

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

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.051*****	-0.009	**0.051*****	**0.026*****	**0.046*****	**0.129*****	**0.215*****	**0.071*****
	(0.010)	(0.010)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Years of Experience	**0.001*****	**0.001*****	**0.001*****	**0.001*****	**0.002*****	**0.002*****	**0.002*****	**0.003*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of Experience missing	-0.023***	-0.017***	-0.025***	-0.029***	-0.033***	-0.032***	-0.027***	-0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.439***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score Missing	0.065***	-0.668***	0.003	-0.782***	-0.571***	0.002	-0.088***	-0.211***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.564***	0.347***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score Missing	-0.343***	0.478***	-0.327***	0.513***	0.360***	-0.107***	0.138***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.193***	-0.197***	-0.186***	-0.283***	-0.279***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.084***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.163***	-0.176***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.054***	-0.057***	-0.067***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.070***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.214***	0.230***	0.147***	0.170***	0.138***	0.079***	0.124***	0.166***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.155***	0.139***	0.159***	0.173***	0.159***	0.162***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.207***	-0.269***	-0.323***	-0.132***	-0.164***	-0.164***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.408***	-0.080**	-0.055	-0.228***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.615***	0.400***	0.304***	0.324***	0.334***	0.356***	0.369***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.616***	-0.429***	-0.541***	-0.588***	-0.246**	-0.360***	-0.200*	-0.291***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.222***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.145***	-0.245***	-0.317***	-0.359***	-0.497***	0.085	-0.214**	-0.298***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.169***	-0.168***	-0.129***	-0.120***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.628***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.282*	-0.823***	-0.462***	-0.605***	-0.249***	-0.407***	-0.672***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.075***	0.009***	-0.015***	-0.021***	-0.004	-0.166***	-0.040***	-0.029***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.534	0.502	0.658	0.605	0.595	0.429	0.416	0.430
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.5

Impact of TFA on achievement; elementary, middle, and high school; controlling for veteran teacher status (>3 yrs expr)

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.058*****	-0.003	**0.057*****	**0.029*****	**0.048*****	**0.143*****	**0.229*****	**0.092*****
	(0.010)	(0.011)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Veteran Teacher (>3 years experience)	**0.038*****	**0.035*****	**0.026*****	**0.015*****	**0.053*****	**0.093*****	**0.090*****	**0.089*****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Teacher experience missing	-0.009***	-0.005***	-0.015***	-0.025***	-0.021***	0.012***	0.017***	0.016***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.439***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score Missing	0.065***	-0.668***	0.003	-0.782***	-0.570***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.575***	0.132***	0.168***	0.564***	0.347***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score Missing	-0.343***	0.478***	-0.327***	0.513***	0.360***	-0.107***	0.138***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.193***	-0.197***	-0.186***	-0.282***	-0.279***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.084***	-0.045***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.164***	-0.088***	-0.102***	-0.142***	-0.112***	-0.163***	-0.176***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.084***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.067***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.070***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.214***	0.230***	0.147***	0.170***	0.138***	0.079***	0.124***	0.166***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.155***	0.139***	0.159***	0.174***	0.159***	0.162***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.207***	-0.269***	-0.323***	-0.132***	-0.164***	-0.164***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.408***	-0.080**	-0.056	-0.226***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.615***	0.400***	0.304***	0.324***	0.334***	0.355***	0.368***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.616***	-0.428***	-0.541***	-0.588***	-0.247**	-0.359***	-0.198*	-0.291***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.222***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.145***	-0.245***	-0.317***	-0.359***	-0.496***	0.084	-0.217**	-0.296***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.169***	-0.168***	-0.129***	-0.120***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.628***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.282*	-0.824***	-0.462***	-0.605***	-0.249***	-0.408***	-0.672***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.061***	-0.002	-0.026***	-0.025***	-0.016***	-0.209***	-0.084***	-0.062***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
R-squared	0.534	0.502	0.658	0.605	0.595	0.429	0.416	0.430
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.6

Impact of TFA on achievement; elementary, middle, and high school; controlling for being fully certified

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.074*****	0.004	**0.066*****	**0.038*****	**0.062*****	**0.177*****	**0.256*****	**0.115*****
	(0.010)	(0.011)	(0.006)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Fully Certified	**0.066*****	**0.045*****	**0.042*****	**0.032*****	**0.091*****	**0.152*****	**0.124*****	**0.151*****
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Certification Status Missing	0.030***	0.025***	0.019***	0.014***	0.056***	0.079***	0.064***	0.098***
	(0.003)	(0.004)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score Missing	0.065***	-0.669***	0.003	-0.784***	-0.573***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.565***	0.347***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score Missing	-0.343***	0.478***	-0.327***	0.514***	0.361***	-0.106***	0.138***	0.342***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.197***	-0.186***	-0.283***	-0.279***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.163***	-0.176***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.084***	-0.048***	-0.031***	-0.028***	-0.054***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.070***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.214***	0.230***	0.147***	0.170***	0.139***	0.079***	0.124***	0.166***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.154***	0.139***	0.160***	0.173***	0.160***	0.163***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.326***	-0.133***	-0.165***	-0.164***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.409***	-0.070*	-0.052	-0.223***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.615***	0.400***	0.303***	0.325***	0.335***	0.356***	0.369***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.619***	-0.429***	-0.542***	-0.591***	-0.244**	-0.362***	-0.205*	-0.289***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.142***	-0.244***	-0.317***	-0.359***	-0.499***	0.087	-0.215**	-0.297***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.170***	-0.168***	-0.129***	-0.121***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.629***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.278*	-0.823***	-0.463***	-0.605***	-0.247***	-0.406***	-0.673***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.016***	-0.029***	-0.056***	-0.057***	-0.084***	-0.292***	-0.139***	-0.154***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
R-squared	0.534	0.501	0.658	0.605	0.594	0.429	0.416	0.430
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.7

Impact of TFA on achievement; elementary, middle, and high school; controlling for having an MA or more

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.040*****	-0.019	**0.048*****	**0.023*****	**0.034*****	**0.106*****	**0.198*****	**0.052*****
	(0.010)	(0.011)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Has MA or more	**0.006*****	0.002	**0.010*****	0.001	**0.004***	**0.010*****	**0.022*****	**0.026*****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Degree status missing	-0.029***	-0.015***	-0.008***	-0.009***	-0.022***	-0.054***	-0.032***	-0.030***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.565***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score Missing	-0.343***	0.478***	-0.327***	0.514***	0.361***	-0.107***	0.137***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.198***	-0.186***	-0.283***	-0.280***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.045***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.164***	-0.177***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.124***	0.167***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.155***	0.139***	0.159***	0.173***	0.160***	0.163***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.327***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.411***	-0.073*	-0.053	-0.226***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.616***	0.400***	0.304***	0.326***	0.335***	0.356***	0.369***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.617***	-0.430***	-0.539***	-0.591***	-0.245**	-0.362***	-0.199*	-0.287***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.143***	-0.245***	-0.318***	-0.360***	-0.502***	0.088	-0.215**	-0.299***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.170***	-0.168***	-0.129***	-0.121***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.629***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.278*	-0.823***	-0.462***	-0.604***	-0.248***	-0.408***	-0.672***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.078***	0.015***	-0.021***	-0.028***	-0.001	-0.152***	-0.036***	-0.018***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.533	0.501	0.658	0.605	0.594	0.428	0.416	0.429
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.8

Impact of TFA on achievement; elementary, middle, and high school; controlling for undergraduate selectivity

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.038*****	-0.020	**0.044*****	**0.023*****	**0.034*****	**0.104*****	**0.190*****	**0.046*****
	(0.010)	(0.010)	(0.005)	(0.006)	(0.010)	(0.009)	(0.011)	(0.012)
Barron's - Most Competitive	**0.100*****	**0.034***	**0.074*****	0.031	-0.108	**-0.070****	**0.454*****	**0.058***
	(0.016)	(0.017)	(0.021)	(0.019)	(0.143)	(0.024)	(0.025)	(0.026)
Barron's - Highly Competitive Plus	-0.044	-0.073	-0.047	**-0.051***	0.020	**-0.186*****	**0.093*****	-0.003
	(0.057)	(0.059)	(0.029)	(0.025)	(0.023)	(0.052)	(0.028)	(0.023)
Barron's - Highly Competitive	**0.037***	0.022	**0.042*****	0.004	**-0.046***	-0.032	0.000	**-0.132*****
	(0.014)	(0.015)	(0.009)	(0.010)	(0.021)	(0.016)	(0.022)	(0.013)
Barron's - Very Competitive Plus	-0.056	-0.054	0.062**	0.026	0.021	**-0.230****	**-0.143****	0.055
	(0.043)	(0.041)	(0.019)	(0.025)	(0.046)	(0.080)	(0.050)	(0.041)
Undergraduate institution missing	-0.030***	-0.016***	-0.011***	-0.009***	-0.024***	-0.057***	-0.040***	-0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.565***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score Missing	-0.343***	0.478***	-0.327***	0.514***	0.361***	-0.107***	0.137***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.198***	-0.186***	-0.283***	-0.280***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.164***	-0.177***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.027***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade school	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.124***	0.167***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.155***	0.139***	0.159***	0.173***	0.160***	0.163***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.326***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.411***	-0.073*	-0.054	-0.226***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.712***	0.616***	0.400***	0.304***	0.326***	0.335***	0.356***	0.370***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.617***	-0.429***	-0.539***	-0.591***	-0.246**	-0.362***	-0.200*	-0.288***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.144***	-0.245***	-0.318***	-0.361***	-0.501***	0.087	-0.214**	-0.299***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.170***	-0.168***	-0.129***	-0.121***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.8 continued

Impact of TFA on achievement; elementary, middle, and high school; controlling for undergraduate selectivity

	Elementary Math	**Elementary Reading**	**Middle School Math**	**Middle School Reading**	**High School English**	**High School Math**	**High School Science**	**High School Social Studies**
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.629***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.278*	-0.823***	-0.461***	-0.604***	-0.248***	-0.407***	-0.671***	-0.476***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.080***	0.015***	-0.018***	-0.028***	0.001	-0.150***	-0.028***	-0.010***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.533	0.501	0.658	0.605	0.594	0.428	0.416	0.429
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.9

Impact of TFA in elementary, middle, and high school; with school-grade-year fixed effects; relative to veteran teachers only

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.036*****	**-0.025***	**0.041*****	**0.025*****	**0.033****	**0.114*****	**0.212*****	**0.040****
	(0.011)	(0.011)	(0.006)	(0.006)	(0.010)	(0.009)	(0.011)	(0.013)
Lagged/8th Grade Reading Score	0.023***	0.455***	0.094***	0.527***	0.428***	0.056***	0.248***	0.378***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged/8th Grade Reading Reading Score	0.054***	-0.693***	-0.007	-0.863***	-0.656***	0.182***	-0.096***	-0.254***
	(0.015)	(0.021)	(0.008)	(0.010)	(0.020)	(0.021)	(0.026)	(0.026)
Lagged/8th Grade Reading Math Score	0.513***	0.082***	0.569***	0.133***	0.164***	0.537***	0.339***	0.181***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged/8th Grade Reading Math Score is	-0.472***	0.366***	-0.356***	0.541***	0.425***	-0.311***	0.135***	0.360***
	(0.016)	(0.022)	(0.008)	(0.010)	(0.020)	(0.021)	(0.026)	(0.026)
Black	-0.298***	-0.279***	-0.187***	-0.209***	-0.236***	-0.211***	-0.310***	-0.319***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Hispanic	-0.063***	-0.105***	-0.067***	-0.078***	-0.115***	-0.079***	-0.119***	-0.109***
	(0.005)	(0.006)	(0.003)	(0.004)	(0.007)	(0.005)	(0.006)	(0.007)
Asian	0.115***	-0.036***	0.185***	-0.016***	-0.030***	0.191***	0.092***	-0.022**
	(0.007)	(0.008)	(0.004)	(0.004)	(0.008)	(0.006)	(0.007)	(0.007)
American Indian	-0.180***	-0.184***	-0.119***	-0.096***	-0.134***	-0.120***	-0.176***	-0.185***
	(0.012)	(0.013)	(0.008)	(0.008)	(0.017)	(0.012)	(0.015)	(0.015)
Other Race	-0.116***	-0.086***	-0.057***	-0.035***	-0.057***	-0.088***	-0.071***	-0.122***
	(0.006)	(0.006)	(0.004)	(0.004)	(0.008)	(0.006)	(0.008)	(0.008)
Female	-0.059***	0.053***	-0.028***	0.044***	0.130***	-0.022***	-0.125***	-0.185***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed: HS or less	0.095***	0.116***	0.067***	0.100***	0.078***	0.033***	0.084***	0.111***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Parent Ed: Some college/trade school	0.219***	0.235***	0.149***	0.177***	0.146***	0.079***	0.138***	0.176***
	(0.003)	(0.004)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Parent Ed: College or more	0.132***	0.147***	0.138***	0.148***	0.143***	0.136***	0.136***	0.166***
	(0.010)	(0.011)	(0.003)	(0.003)	(0.006)	(0.005)	(0.006)	(0.006)
Student has disability	-0.325***	-0.400***	-0.215***	-0.275***	-0.337***	-0.145***	-0.166***	-0.176***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Disability status missing	0.000	0.000	0.000	0.000	-0.447***	-0.093	0.035	-0.258***
	(.)	(.)	(.)	(.)	(0.077)	(0.062)	(0.067)	(0.058)
Student is gifted	0.704***	0.626***	0.389***	0.294***	0.315***	0.323***	0.343***	0.349***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)
Gifted status missing	-0.808***	-0.339**	-0.531***	-0.697***	-0.308*	-0.347***	-0.130	-0.166
	(0.079)	(0.128)	(0.050)	(0.087)	(0.153)	(0.089)	(0.172)	(0.118)
Student ever IDed as LEP	-0.128***	-0.239***	-0.040***	-0.163***	-0.264***	-0.014**	-0.127***	-0.184***
	(0.005)	(0.006)	(0.003)	(0.004)	(0.007)	(0.005)	(0.006)	(0.006)
LEP status missing	-0.122	-0.225**	-0.238***	-0.408***	-0.536***	0.172*	-0.452***	-0.395***
	(0.066)	(0.073)	(0.042)	(0.050)	(0.107)	(0.073)	(0.131)	(0.095)
Student over age for grade	-0.168***	-0.170***	-0.131***	-0.123***	-0.213***	-0.193***	-0.247***	-0.261***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student under age for grade	0.054***	0.064***	0.056***	0.031***	0.104***	0.144***	0.181***	0.157***
	(0.010)	(0.011)	(0.005)	(0.006)	(0.010)	(0.007)	(0.008)	(0.009)
Student has repeated grade	-0.879***	-0.951***	-0.406***	-0.438***	-0.394***	-0.504***	-0.598***	-0.582***
	(0.007)	(0.008)	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Repeater status missing	-1.168**	-1.008**	-0.324**	-0.470**	-0.251***	-0.409***	-0.692***	-0.332***
	(0.452)	(0.336)	(0.123)	(0.150)	(0.038)	(0.035)	(0.059)	(0.048)
Constant	0.100***	0.023***	-0.015***	-0.022***	0.051***	-0.173***	-0.055***	0.042***
	(0.005)	(0.005)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
R-squared	0.529	0.501	0.664	0.613	0.597	0.420	0.416	0.436
Observations	449240	445643	910910	906295	271053	599659	473207	403987

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Grade 6 is included in middle school as 90% of students in NC schools change schools between grades 5 and 6.

Veteran teachers have 4 or more years of experience.

Appendix Table 4.10

Impact of TFA in elementary, middle, and high school; school-grade-year fixed effects; relative to veteran teachers only; testing for differences in early and late TFA

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	0.013	**-0.084***	0.011	0.009	0.021	**0.157*****	**0.150*****	**-0.470*****
	(0.040)	(0.042)	(0.012)	(0.013)	(0.023)	(0.032)	(0.022)	(0.072)
TFA*Post-2005	0.025	0.063	**0.040****	0.021	0.016	-0.048	**0.084*****	**0.527*****
	(0.042)	(0.044)	(0.014)	(0.015)	(0.025)	(0.033)	(0.025)	(0.073)
Lagged/8th grade reading score	0.023***	0.455***	0.094***	0.527***	0.428***	0.056***	0.248***	0.378***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged/8th grade reading score is missing	0.054***	-0.693***	-0.007	-0.863***	-0.656***	0.182***	-0.096***	-0.254***
	(0.015)	(0.021)	(0.008)	(0.010)	(0.020)	(0.021)	(0.026)	(0.026)
Lagged/8th grade math score	0.513***	0.082***	0.569***	0.133***	0.164***	0.537***	0.339***	0.181***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged/8th grade math score is missing	-0.472***	0.366***	-0.356***	0.541***	0.425***	-0.311***	0.135***	0.360***
	(0.016)	(0.022)	(0.008)	(0.010)	(0.020)	(0.021)	(0.026)	(0.026)
Black	-0.298***	-0.279***	-0.187***	-0.209***	-0.236***	-0.211***	-0.310***	-0.319***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Hispanic	-0.063***	-0.105***	-0.067***	-0.078***	-0.115***	-0.079***	-0.119***	-0.109***
	(0.005)	(0.006)	(0.003)	(0.004)	(0.007)	(0.005)	(0.006)	(0.007)
Asian	0.115***	-0.036***	0.185***	-0.016***	-0.030***	0.191***	0.092***	-0.022**
	(0.007)	(0.008)	(0.004)	(0.004)	(0.008)	(0.006)	(0.007)	(0.007)
American Indian	-0.180***	-0.184***	-0.119***	-0.096***	-0.134***	-0.120***	-0.176***	-0.185***
	(0.012)	(0.013)	(0.008)	(0.008)	(0.017)	(0.012)	(0.015)	(0.015)
Other Race	-0.116***	-0.086***	-0.057***	-0.035***	-0.057***	-0.088***	-0.071***	-0.122***
	(0.006)	(0.006)	(0.004)	(0.004)	(0.008)	(0.006)	(0.008)	(0.008)
Female	-0.059***	0.053***	-0.028***	0.044***	0.130***	-0.022***	-0.125***	-0.185***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Parent Ed: HS or less	0.095***	0.116***	0.067***	0.100***	0.078***	0.033***	0.084***	0.111***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Parent Ed: Some college/trade school	0.219***	0.235***	0.149***	0.177***	0.146***	0.079***	0.138***	0.176***
	(0.003)	(0.004)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Parent Ed: College or more	0.132***	0.147***	0.138***	0.148***	0.143***	0.136***	0.136***	0.166***
	(0.010)	(0.011)	(0.003)	(0.003)	(0.006)	(0.005)	(0.006)	(0.006)
Student has disability	-0.325***	-0.400***	-0.215***	-0.275***	-0.337***	-0.145***	-0.166***	-0.176***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Disability status missing	0.000	0.000	0.000	0.000	-0.447***	-0.093	0.035	-0.259***
	(.)	(.)	(.)	(.)	(0.077)	(0.062)	(0.067)	(0.058)
Student is gifted	0.704***	0.626***	0.389***	0.294***	0.314***	0.323***	0.343***	0.349***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)
Gifted status missing	-0.808***	-0.339**	-0.531***	-0.697***	-0.308*	-0.347***	-0.130	-0.166
	(0.079)	(0.128)	(0.050)	(0.087)	(0.153)	(0.089)	(0.172)	(0.118)
Student ever IDed as LEP	-0.128***	-0.239***	-0.040***	-0.163***	-0.264***	-0.014**	-0.127***	-0.184***
	(0.005)	(0.006)	(0.003)	(0.004)	(0.007)	(0.005)	(0.006)	(0.006)
LEP status missing	-0.122	-0.225**	-0.238***	-0.408***	-0.536***	0.172*	-0.452***	-0.395***
	(0.066)	(0.073)	(0.042)	(0.050)	(0.107)	(0.073)	(0.131)	(0.095)
Student over age for grade	-0.168***	-0.170***	-0.131***	-0.123***	-0.213***	-0.193***	-0.247***	-0.261***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student under age for grade	0.054***	0.064***	0.056***	0.031***	0.104***	0.144***	0.181***	0.157***
	(0.010)	(0.011)	(0.005)	(0.006)	(0.010)	(0.007)	(0.008)	(0.009)
Student has repeated grade	-0.879***	-0.951***	-0.406***	-0.438***	-0.394***	-0.504***	-0.598***	-0.582***
	(0.007)	(0.008)	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Repeater status missing	-1.168**	-1.008**	-0.323**	-0.470**	-0.251***	-0.409***	-0.692***	-0.332***
	(0.452)	(0.336)	(0.123)	(0.150)	(0.038)	(0.035)	(0.059)	(0.048)
Constant	0.100***	0.023***	-0.015***	-0.022***	0.051***	-0.173***	-0.055***	0.041***
	(0.005)	(0.005)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
R-squared	0.529	0.501	0.664	0.613	0.597	0.420	0.417	0.436
Observations	449240	445643	910910	906295	271053	599659	473207	403987

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Grade 6 is included in middle school as 90% of students in NC schools change schools between grades 5 and 6.

Veteran teachers have 4 or more years of experience.

Appendix Table 4.11

Impact of TFA in elementary and middle school, using school-grade-year fixed effects, testing whether teacher quality can explain time-trends in TFA effectiveness; Quality measure: Mean Praxis scores

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	0.015	**-0.082***	0.019	0.009	0.012	**0.176*****	**0.130*****	**-0.440*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.014	0.064	**0.028***	0.017	0.020	**-0.088****	**0.071****	**0.482*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Mean Praxis Score	**0.014*****	**0.004*****	**0.008*****	**0.004*****	**0.022*****	**0.019*****	**0.034*****	**0.043*****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean Praxis Score*Post-2005	**0.004*****	-0.001	**0.003*****	**-0.002***	**-0.015*****	**-0.012*****	**-0.017*****	**-0.017*****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Praxis score missing	-0.027***	-0.014***	-0.016***	-0.011***	-0.025***	-0.065***	-0.046***	-0.043***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Lagged/8th Grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Reading Score is Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.001	-0.088***	-0.211***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th Grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.565***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th Grade Math Score is Missing	-0.343***	0.478***	-0.327***	0.514***	0.361***	-0.106***	0.138***	0.340***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.197***	-0.186***	-0.283***	-0.279***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.141***	-0.112***	-0.164***	-0.176***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.099***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.124***	0.166***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Missing	0.148***	0.155***	0.139***	0.159***	0.173***	0.160***	0.163***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.320***	-0.397***	-0.208***	-0.271***	-0.326***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status is missing	0.000	0.000	0.000	0.000	-0.411***	-0.075*	-0.055	-0.225***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.615***	0.400***	0.304***	0.325***	0.335***	0.356***	0.369***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.617***	-0.429***	-0.540***	-0.591***	-0.246**	-0.360***	-0.202*	-0.289***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.144***	-0.245***	-0.319***	-0.360***	-0.501***	0.085	-0.215**	-0.300***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.169***	-0.168***	-0.129***	-0.121***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.630***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.279*	-0.824***	-0.461***	-0.605***	-0.249***	-0.407***	-0.674***	-0.476***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.080***	0.016***	-0.016***	-0.027***	0.003	-0.143***	-0.024***	-0.010***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.534	0.501	0.658	0.605	0.594	0.429	0.416	0.430
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Grade 6 is included in middle school as 90% of students in NC schools change schools between grades 5 and 6.

Veteran teachers have 4 or more years of experience.

Appendix Table 4.12

Impact of TFA in elementary, middle, and high school; using school-grade-year fixed effects; testing whether teacher quality can explain time-trends in TFA effectiveness; Quality measure: Years of Experience

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	0.020	-0.077	0.019	0.009	0.032	**0.173*****	**0.167*****	**-0.423*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.032	0.072	**0.041****	0.021	0.016	-0.052	**0.060***	**0.502*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Years of Experience	**0.002*****	**0.001*****	**0.001*****	**0.001*****	**0.003*****	**0.003*****	**0.003*****	**0.005*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of Experience*Post-2005	**-0.001*****	**-0.000***	-0.000	**-0.001*****	**-0.002*****	**-0.002*****	**-0.003*****	**-0.003*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of Experience missing	-0.023***	-0.017***	-0.026***	-0.029***	-0.033***	-0.031***	-0.026***	-0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.439***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Reading Score is Missing	0.065***	-0.668***	0.003	-0.782***	-0.571***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.564***	0.347***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Math Score is Missing	-0.344***	0.478***	-0.327***	0.513***	0.361***	-0.107***	0.138***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.193***	-0.197***	-0.186***	-0.283***	-0.279***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.084***	-0.045***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.163***	-0.176***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.054***	-0.057***	-0.067***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.099***	0.117***	0.070***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.214***	0.230***	0.147***	0.170***	0.138***	0.079***	0.124***	0.166***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Missing	0.148***	0.154***	0.139***	0.159***	0.173***	0.159***	0.162***	0.193***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.207***	-0.269***	-0.323***	-0.132***	-0.163***	-0.163***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status is missing	0.000	0.000	0.000	0.000	-0.410***	-0.082**	-0.059	-0.230***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.615***	0.400***	0.304***	0.324***	0.334***	0.356***	0.368***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.616***	-0.429***	-0.541***	-0.589***	-0.245**	-0.359***	-0.197*	-0.290***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.222***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.145***	-0.245***	-0.317***	-0.359***	-0.499***	0.084	-0.216**	-0.299***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.169***	-0.168***	-0.129***	-0.120***	-0.201***	-0.188***	-0.236***	-0.243***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.629***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.282*	-0.824***	-0.462***	-0.605***	-0.249***	-0.408***	-0.671***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.076***	0.009***	-0.015***	-0.021***	-0.003	-0.166***	-0.039***	-0.028***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.534	0.502	0.658	0.605	0.595	0.429	0.416	0.430
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.13

Impact of TFA in high school, using school-grade-year fixed effects, testing whether teacher quality can explain time-trends in TFA effectiveness; Quality measure: Veteran teacher

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	0.027	-0.072	**0.026***	0.011	0.034	**0.174*****	**0.173*****	**-0.392*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.030	0.072	**0.040****	0.020	0.014	-0.041	**0.069****	**0.488*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Veteran teacher (>3 years experience)	**0.045*****	**0.038*****	**0.028*****	**0.020*****	**0.078*****	**0.115*****	**0.116*****	**0.123*****
	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
Veteran teacher*Post-2005	**-0.015*****	**-0.006****	**-0.003***	**-0.013*****	**-0.049*****	**-0.042*****	**-0.054*****	**-0.057*****
	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Years of Experience missing	-0.009***	-0.005***	-0.015***	-0.025***	-0.020***	0.014***	0.019***	0.014***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.439***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Reading Score is Missing	0.065***	-0.668***	0.003	-0.782***	-0.571***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th grade Math Score	0.497***	0.070***	0.575***	0.132***	0.168***	0.564***	0.347***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Math Score is Missing	-0.343***	0.478***	-0.327***	0.513***	0.361***	-0.106***	0.138***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.193***	-0.197***	-0.186***	-0.282***	-0.279***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.084***	-0.045***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.163***	-0.176***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.084***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.067***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: Some college/trade school	0.099***	0.117***	0.070***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.214***	0.230***	0.147***	0.170***	0.138***	0.079***	0.124***	0.166***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Missing	0.148***	0.155***	0.139***	0.159***	0.173***	0.158***	0.162***	0.193***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.207***	-0.269***	-0.323***	-0.132***	-0.163***	-0.163***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status is missing	0.000	0.000	0.000	0.000	-0.411***	-0.083**	-0.059	-0.230***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.615***	0.400***	0.304***	0.324***	0.334***	0.355***	0.368***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.616***	-0.428***	-0.541***	-0.589***	-0.245**	-0.357***	-0.196*	-0.290***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.222***	-0.000	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.145***	-0.245***	-0.317***	-0.359***	-0.499***	0.083	-0.218**	-0.297***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.169***	-0.168***	-0.129***	-0.120***	-0.201***	-0.188***	-0.236***	-0.243***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.628***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.282*	-0.824***	-0.462***	-0.605***	-0.250***	-0.408***	-0.672***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.061***	-0.002	-0.026***	-0.025***	-0.016***	-0.210***	-0.085***	-0.060***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
R-squared	0.534	0.502	0.658	0.605	0.595	0.429	0.416	0.430
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.14

Impact of TFA on same-year tests; elementary, middle, and high school; school-grade-year fixed effects; by early and established TFA; controlling for teacher quality; teacher quality measure: Fully certified

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	0.053	-0.055	**0.042*****	**0.026***	**0.075*****	**0.233*****	**0.211*****	**-0.339*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.015	0.055	**0.030***	0.015	-0.018	**-0.068***	**0.060***	**0.470*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Fully Certified	**0.074*****	**0.053*****	**0.045*****	**0.034*****	**0.110*****	**0.175*****	**0.132*****	**0.146*****
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
Fully Certified*Post-2005	**-0.025*****	**-0.023*****	**-0.007*****	**-0.005***	**-0.035*****	**-0.042*****	**-0.015*****	**0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Certification Status Missing	0.029***	0.025***	0.020***	0.014***	0.063***	0.087***	0.067***	0.096***
	(0.003)	(0.004)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Reading Score is Missing	0.065***	-0.669***	0.003	-0.784***	-0.573***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.565***	0.347***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Math Score is Missing	-0.343***	0.478***	-0.327***	0.514***	0.361***	-0.106***	0.138***	0.342***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.197***	-0.186***	-0.283***	-0.279***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.163***	-0.176***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.054***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.070***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade sch	0.214***	0.230***	0.147***	0.170***	0.139***	0.079***	0.124***	0.166***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.154***	0.139***	0.159***	0.172***	0.159***	0.162***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.326***	-0.133***	-0.165***	-0.164***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.415***	-0.078**	-0.055	-0.221***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.711***	0.615***	0.400***	0.303***	0.325***	0.334***	0.356***	0.369***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.617***	-0.429***	-0.541***	-0.591***	-0.242**	-0.357***	-0.203*	-0.289***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.144***	-0.246***	-0.318***	-0.360***	-0.503***	0.083	-0.216**	-0.297***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.14 continued

Impact of TFA on same-year tests; elementary, middle, and high school; school-grade-year fixed effects; by early and established TFA; controlling for teacher quality; teacher quality measure: Fully certified

	Elementary Math	**Elementary Reading**	**Middle School Math**	**Middle School Reading**	**High School English**	**High School Math**	**High School Science**	**High School Social Studies**
Student over age for grade	-0.169***	-0.168***	-0.129***	-0.121***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.629***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.278*	-0.823***	-0.463***	-0.605***	-0.247***	-0.408***	-0.673***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.021***	-0.024***	-0.055***	-0.057***	-0.084***	-0.292***	-0.139***	-0.153***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
R-squared	0.534	0.501	0.658	0.605	0.594	0.429	0.416	0.430
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.15

Impact of TFA on same-year tests; elementary, middle, and high school; school-grade-year fixed effects; by early and established TFA; controlling for teacher quality; teacher quality measure: Has an MA or more

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	0.003	**-0.090***	0.016	0.005	0.007	**0.133*****	**0.122*****	**-0.475*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.040	0.074	**0.042****	0.022	0.035	-0.030	**0.100*****	**0.542*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Has MA or more	**0.008*****	**0.006*****	**0.008*****	**0.005*****	**0.011*****	**0.019*****	**0.035*****	**0.034*****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
Has MA or more*Post-2005	**-0.004***	**-0.007*****	**0.005****	**-0.008*****	**-0.013*****	**-0.015*****	**-0.023*****	**-0.013*****
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Degree status missing	-0.029***	-0.015***	-0.008***	-0.009***	-0.021***	-0.052***	-0.030***	-0.029***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.252***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Reading Score is Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.002	-0.088***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.565***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Math Score is Missing	-0.343***	0.478***	-0.327***	0.514***	0.361***	-0.107***	0.138***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.198***	-0.186***	-0.283***	-0.280***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.045***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.164***	-0.177***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.028***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade sch	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.124***	0.167***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.154***	0.139***	0.159***	0.173***	0.160***	0.163***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.327***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.711***	0.616***	0.400***	0.304***	-0.411***	-0.074*	-0.055	-0.226***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.000	0.000	0.000	0.000	0.326***	0.335***	0.356***	0.369***
	(.)	(.)	(.)	(.)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.617***	-0.430***	-0.539***	-0.591***	-0.245**	-0.362***	-0.199*	-0.287***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.143***	-0.245***	-0.318***	-0.361***	-0.503***	0.087	-0.215**	-0.299***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.170***	-0.168***	-0.129***	-0.121***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.629***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.278*	-0.823***	-0.462***	-0.604***	-0.248***	-0.408***	-0.672***	-0.477***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.078***	0.015***	-0.021***	-0.028***	-0.001	-0.153***	-0.037***	-0.018***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.533	0.501	0.658	0.605	0.594	0.428	0.416	0.429
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.16

Impact of TFA on same-year tests; elementary, middle, and high school; school-grade-year fixed effects; by early and established TFA; controlling for teacher quality; teacher quality measure: Barron's selectivity rankings of undergraduate institution

	Elementary Math	Elementary Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	0.001	**-0.092***	0.014	0.005	0.005	**0.129*****	**0.122*****	**-0.480*****
	(0.038)	(0.040)	(0.011)	(0.013)	(0.021)	(0.032)	(0.021)	(0.071)
TFA*Post-2005	0.040	0.077	**0.039****	0.023	0.037	-0.027	**0.095*****	**0.542*****
	(0.040)	(0.042)	(0.013)	(0.014)	(0.023)	(0.033)	(0.024)	(0.072)
Barron's - Most Competitive	**0.088*****	**0.044***	**0.072****	0.039	-0.108	**-0.073***	**0.494*****	-0.007
	(0.020)	(0.021)	(0.023)	(0.025)	(0.143)	(0.033)	(0.029)	(0.043)
Barron's - Highly Competitive Plus	**-1.665****	-0.073	-0.076	-0.029	**0.205*****	**-0.470*****	**0.151*****	**-0.213****
	(0.633)	(0.059)	(0.041)	(0.045)	(0.041)	(0.096)	(0.039)	(0.069)
Barron's - Highly Competitive	0.009	0.038	-0.002	-0.019	0.003	-0.024	**-0.179*****	-0.027
	(0.028)	(0.029)	(0.016)	(0.017)	(0.033)	(0.023)	(0.033)	(0.029)
Barron's - Very Competitive Plus	-0.081	-0.099	0.043	0.007	0.021	**-0.230****	-0.140	0.056
	(0.055)	(0.059)	(0.025)	(0.027)	(0.046)	(0.080)	(0.097)	(0.041)
Barron's - Most Competitive*Post-2005	0.035	-0.026	0.013	-0.020	0.000	0.007	**-0.165****	0.102
	(0.034)	(0.034)	(0.060)	(0.039)	(.)	(0.049)	(0.058)	(0.053)
Barron's - Highly Comp. Plus*Post-2005	**1.635***	0.000	0.058	-0.031	**-0.277*****	**0.403*****	**-0.122***	**0.236****
	(0.635)	(.)	(0.058)	(0.054)	(0.050)	(0.115)	(0.055)	(0.073)
Barron's - Highly Comp.*Post-2005	0.039	-0.021	**0.063*****	0.036	-0.079	-0.016	**0.339*****	**-0.134*****
	(0.033)	(0.034)	(0.019)	(0.022)	(0.042)	(0.033)	(0.045)	(0.032)
Barron's - Very Comp. Plus*Post-2005	0.061	0.089	0.050	0.120	0.000	0.000	0.006	0.000
	(0.086)	(0.082)	(0.040)	(0.067)	(.)	(.)	(0.113)	(.)
Undergraduate institution missing	-0.030***	-0.016***	-0.011***	-0.009***	-0.024***	-0.057***	-0.040***	-0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Lagged/8th grade Reading Score	0.022***	0.451***	0.092***	0.529***	0.440***	0.061***	0.253***	0.389***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Reading Score is Missing	0.064***	-0.669***	0.002	-0.784***	-0.573***	0.002	-0.087***	-0.212***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Lagged/8th grade Math Score	0.497***	0.070***	0.576***	0.132***	0.168***	0.565***	0.348***	0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged/8th grade Math Score is Missing	-0.343***	0.478***	-0.327***	0.514***	0.361***	-0.107***	0.137***	0.341***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.009)	(0.009)	(0.011)	(0.010)
Black	-0.309***	-0.288***	-0.163***	-0.194***	-0.198***	-0.186***	-0.283***	-0.280***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.068***	-0.099***	-0.038***	-0.059***	-0.085***	-0.044***	-0.084***	-0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Asian	0.117***	-0.037***	0.160***	-0.018***	-0.025***	0.208***	0.101***	0.013***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
American Indian	-0.151***	-0.165***	-0.088***	-0.102***	-0.142***	-0.112***	-0.164***	-0.177***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.006)
Other Race	-0.105***	-0.085***	-0.048***	-0.031***	-0.028***	-0.055***	-0.057***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Female	-0.070***	0.045***	-0.027***	0.042***	0.137***	-0.020***	-0.142***	-0.202***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parent Ed: HS or less	0.099***	0.117***	0.071***	0.101***	0.080***	0.036***	0.082***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: Some college/trade sch	0.214***	0.230***	0.147***	0.170***	0.139***	0.080***	0.124***	0.167***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Parent Ed: College or more	0.148***	0.154***	0.139***	0.159***	0.173***	0.160***	0.163***	0.194***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix Table 4.16 continued

Impact of TFA on same-year tests; elementary, middle, and high school; school-grade-year fixed effects; by early and established TFA; controlling for teacher quality; teacher quality measure: Barron's selectivity rankings of undergraduate institution

Student has disability	-0.319***	-0.397***	-0.208***	-0.271***	-0.327***	-0.133***	-0.165***	-0.165***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Disability status missing	0.000	0.000	0.000	0.000	-0.411***	-0.073*	-0.054	-0.226***
	(.)	(.)	(.)	(.)	(0.032)	(0.029)	(0.033)	(0.029)
Student is gifted	0.712***	0.616***	0.400***	0.304***	0.326***	0.335***	0.356***	0.370***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Gifted status missing	-0.617***	-0.429***	-0.539***	-0.591***	-0.246**	-0.362***	-0.199*	-0.288***
	(0.031)	(0.048)	(0.030)	(0.051)	(0.083)	(0.057)	(0.099)	(0.072)
Student ever IDed as LEP	-0.124***	-0.252***	-0.032***	-0.152***	-0.223***	-0.001	-0.116***	-0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
LEP status missing	-0.144***	-0.245***	-0.318***	-0.360***	-0.501***	0.087	-0.214**	-0.298***
	(0.025)	(0.027)	(0.025)	(0.028)	(0.060)	(0.046)	(0.069)	(0.053)
Student over age for grade	-0.170***	-0.168***	-0.129***	-0.121***	-0.201***	-0.188***	-0.236***	-0.244***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student under age for grade	0.061***	0.049***	0.065***	0.035***	0.109***	0.151***	0.190***	0.165***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)
Student has repeated grade	-0.892***	-0.978***	-0.421***	-0.467***	-0.421***	-0.513***	-0.634***	-0.629***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)
Repeater status missing	-0.278*	-0.823***	-0.460***	-0.604***	-0.248***	-0.407***	-0.671***	-0.476***
	(0.135)	(0.153)	(0.063)	(0.073)	(0.022)	(0.021)	(0.032)	(0.028)
Constant	0.080***	0.015***	-0.018***	-0.028***	0.001	-0.150***	-0.028***	-0.010***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.533	0.501	0.658	0.605	0.594	0.428	0.416	0.429
Observations	3793320	3770563	3734293	3722294	1270188	2622518	2174625	1865239

* p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors in parentheses.

Appendix A

Methodological Supplement

This methodological supplement details several technical issues regarding weighting, variable choice, missing data, inconsistent scoring metrics across grades, and comparisons to another distributional paper examining Teach For America.

Identification of Quantile Treatment Effects

The analyses presented in the accompanying manuscript take advantage of random assignment into Teach For America (TFA) or non-TFA classrooms to estimate the mean effect of having a TFA teacher in elementary school as well as its effect on the distribution of student achievement. The potential outcomes model provides a framework for estimation of the effects of such a treatment. Each individual i has two potential outcomes, Y_{1i} and Y_{0i} (for my purposes, a test score). Person i has outcome Y_{1i} if assigned to the treatment group and outcome Y_{0i} if assigned to the control group. D(i) denotes the group that i is assigned to in a randomized experiment. If person i is assigned to the treatment group, then D(i) = 1, and if person i is assigned to the control group, D(i) = 0; the treatment effect on person i is defined as $d_i=Y_{1i} - Y_{0i}$.

Let Y be a random variable with a cumulative distribution function (CDF) F(y), where $F(y) = Pr[Y \leq y]$. Then, the qth quantile of the distribution F(y) is defined as the smallest value y_q such that $F(y_q)$ is at least as large as q (e.g., $y_{0.5}$ is the median). Now consider two (marginal) distributions F_1 (the CDF for the potential outcomes if D = 1), and F_0 (the CDF for the potential outcomes if D = 0). We define the difference between the qth quantiles of these two distributions as $y_q = y_{q1} - y_{q0}$, where y_{qd} is the qth quantile of distribution F_d.

The joint distribution of (Y_{0i},Y_{1i}) is not identified without assumptions. However, if

program assignment is independent of the potential outcomes, the difference in means, or average treatment effect, $d = E[d_i] = E[Y_1] - E[Y_0]$, is identified because each expectation requires only one of the two marginal distributions. Similarly, identification of the marginal distributions implies identification of the quantiles y_{qd}, and thus identification of the differences in their quantiles, $y_q = y_{q1} - y_{q0}$. In this experimental setting, the quantile treatment effect (QTE) is the estimate of this difference in the quantiles of the two marginal distributions. For example, we consistently estimate the QTE at the 0.50 quantile by subtracting the control group's sample median from the treatment group's sample median. Graphically, QTE estimates are the horizontal differences in the CDFs of the outcome for the treatment and control groups at various percentiles, as shown in Figures 1 and 2 in the manuscript.

Outcome Variable Selection

For my dependent variable, I follow Glazerman et al. and use Normal Curve Equivalent (NCE) scores. NCE scores are calculated from raw scores that are normed based on grade and quarter of the school year (fall, winter, or spring) using the national ITBS sample, and converted into rankings such that the distribution of scores makes a normal distribution. This allows for cross-age and cross-grade comparisons of scores that are equal interval. Because these scores are normed by age and quarter, students who are growing at the average rate should have scores that are no different at the two time points. That is, a student who is at the national median in the fall of second grade would receive a 50, as would a student who is at the median in the spring of fifth grade, even though the student in the spring of fifth grade will know more than the student in the fall of fifth grade. As noted in Glazerman et al., math achievement of students in TFA classrooms increased by 2 NCE points on average, while average achievement for students in non-TFA classrooms stayed constant. In contrast, for reading, the average achievement for

students in TFA and non-TFA classrooms increased by 1 NCE point from fall to spring. NCE scores are also better suited for distributional comparisons than National Percentile Ranking or Number Correct scores because they are on an equal interval scale.

Invalid test scores

In addition to the test scores that were identified as missing in the public use data, 155 of the respondents (including 9.5 percent of respondents in grades 2-5) have at least one NCE test score of 0, which correspond to raw scores of 99. According to the ITBS website for the publisher, Riverside Publishing, and confirmed in a telephone conversation with customer support, students are given tests of increasing difficulty depending on age and skill level in timed sessions that do not exceed 30 minutes. Raw scores are calculated from each test level. Although the total number of questions varies somewhat by level, the highest possible raw score in reading at any level is 44 and the highest possible raw score in math is 50 (Hoover, Dunbar, & Frisbie, 2013). Thus, as noted in the paper, I treat the NCE scores of 0 and raw scores of 99 as invalid test scores. Figure A.1 presents a histogram of the raw scores from the fall math test, showing how far above the other scores these 99s fall.

To my knowledge, no other study has noted this. The degree to which this has been noticed is thus unclear, as it is not mentioned in other work using these data (Antecol, Eren, & Ozbeklik, 2013; Decker et al., 2004b; Glazerman et al., 2006). Fortunately, the prevalence of 99s in posttests (spring tests) is fairly balanced across TFA and non-TFA classrooms for both the full sample and my analytic sample. However, there are statistically significant differences in their prevalence in both the reading and math pretests (fall tests) for the full sample and my analytic sample, with non-TFA classrooms having a higher proportion in reading and TFA classrooms having a higher proportion in math. As shown in Table 1.A, inclusion of the inverse propensity

score weights balances these group differences as well as those associated with valid missings at baseline and in the posttest.

Different Reading Scores for Grade One

Although students in grades 2-5 received reading scores that were calculated using responses from tests for both vocabulary and word analysis, first grade reading tests were scored separately as vocabulary and word analysis. While the public use documentation provided by MPR suggests using one of these scores in place of the total reading score used for children in higher grades, in order to ensure that results across grades were comparable, I use only students in grades two through five. Communication with Glazerman (2013) suggests that the Glazerman et al. (2006) paper used the vocabulary score in place of a full reading score for first graders. Supplemental analyses following Glazerman et al.'s (2006) approach, including first graders and using their vocabulary score, yield similar results to those reported.

Floor and Ceiling Effects

After recoding invalid NCE scores of zero, there is still some clustering at the floor score of one on the fall and spring math and reading tests. However, this concentration is limited to only the lowest few percentiles on the spring tests. In the fall and spring math tests, students with an NCE score of one are in the 6^{th} percentile and below, suggesting a minimal floor effect. Likewise, on the fall and spring reading tests, only students at the 8^{th} percentile and below have the floor score of one, which suggests that they are not driving the negative differences observed below the 32^{nd} percentile in the spring QTE.

Ceiling effects are even less of a concern. On the spring math test, the 98^{th} percentile student in these data has an NCE score of 79 points, and in reading the highest score is 90 NCE

points. The fall tests in both math and reading have even lower high scores. These scores are well below the maximum NCE score.

Inverse Propensity Score Weights

This paper draws on Firpo (2007) and Glazerman, Mayer, & Decker (2006), estimating the inverse propensity score weighted quantile treatment effects (QTE) of being assigned to a Teach For America (TFA) classroom. In doing so, I extend previous research on the effectiveness of TFA beyond examining only mean differences, and consider how it might affect the distribution of student achievement more broadly. I create inverse propensity score weights to adjust for nonresponse (including non-response because the observation is missing a pre-test score as previously identified by MPR and non-response from treating the invalid 99 raw scores as missing data). Inverse propensity score weighting provides unconditional QTE while providing a way to ensure that treatment and control are balanced on observed characteristics. Readers interested in technical details about this approach are encouraged to consult Firpo (2007).

Briefly, I created an inverse propensity score weight using predicted probabilities from a logistic regression model predicting treatment status as a function of randomization block, indicators for missing and invalid test scores (see below), baseline test score deciles, and demographic characteristics (including race and gender). As described in the manuscript, I calculate a predicted probability of being in the treatment group $\hat{p}$, and then construct weights of $1/\hat{p}$ for those in the treatment group and $1/(1-\hat{p})$ for the control group. In deciding what covariates to include in this logistic regression model, I draw heavily on the work of Glazerman et al. (2006).

Inverse propensity score weight adjusted fall QTE, are shown in Figure A.2. Panels A and B. Confidence intervals for the point estimates are calculated by bootstrapping within

randomization block and treatment status 999 times, re-estimating test score deciles and the logistic regression for the inverse propensity scores in each sample.

[Insert Figure A.2 here]

As shown in Figure A.2, Panels A and B, the inverse propensity score weights described above are quite successful at balancing differences between treatment and control across the entire distribution for both math and reading. For math, the confidence intervals always include zero, and for most of the distribution, the point estimates are at or near zero. Likewise, the reading point estimates are near zero for most of the distribution, except at the very upper tail. Although the confidence intervals are fairly wide near the lower tail, the confidence intervals always include zero.[40]

The basic pattern of results reported here are robust to a variety of weighting equations, including weights that make the TFA classrooms look more like the non-TFA classrooms (in other words Treatment on the non-Treated), weights that make the non-TFA classrooms look like the TFA classrooms (Treatment on the Treated), and weights that make TFA and non-TFA classrooms look more like the overall sample (the approach used in the current draft). The results are also robust to the inclusion of several different combinations of demographic controls and indicators for missing data in the logistic regression. In particular, the posttest results are largely consistent when the model includes only randomization block, when the model includes everything but the baseline demographic characteristics, and when the model substitutes linear test score terms in place of the test score deciles. The balance at baseline is not always as successful in these various model specifications, and thus the model described in the paper, which achieves the best level of balance on baseline test scores, was selected as the primary

model. In each model specification, to bootstrap the standard errors, the inverse propensity score weights were recalculated on each new bootstrapped sample.

Robustness Checks

There are several potential mechanisms which might explain some of the observed variation in effects across the distribution of student achievement in reading. One possibility is that a particular TFA region is driving the results, possibly due to variation in language arts curricula or instructional practices. Although the original Mathematica report suggests that the impacts were relatively similar across regions and that, "overall impacts were not attributable to any particular region, school, or grade" (p. xv), it is worth confirming that the same is true in terms of the impact of TFA on the distribution. Because the sample sizes in any one region are small and would produce noisy estimates at best, instead I performed a series of six models in which one region was dropped from the analysis each time. Across the post-test results for each of the six models, the shape of the quantile treatment effect graphs remained qualitatively consistent with the models including all six regions. These robustness models rule out regional and district differences as a potential mechanism driving these patterns of effects. It is not the case that teacher practices or curricula in a particular district lead to vastly different results across the distribution.

A second potential mechanism that could be driving the negative TFA effects at the bottom of the distribution is differences in the language in which the test was administered. Specifically, some students were given the Spanish version of the ITBS, which could contribute to the negative differences at the bottom of the distribution. Perhaps surprisingly, fewer than 10 percent of students are tested in Spanish or are in bilingual classrooms, and many of these students have relatively high scores on both the reading and math posttests. This suggests that the

language of the test administration is also not the cause of the variation in effects. These robustness analyses provide at least some sense for two mechanisms that do not seem to be driving the pattern of effects, although they are not exhaustive tests of potential mechanisms.

Significance testing of results

The results for the reading post-test models provide evidence of a countervailing pattern across the distribution, with negative effects at the bottom of the distribution and positive effects at the top. While there are some portions of the distribution where these differences are significant at the 5% level, it is also important to confirm that these results are indicative of a larger trend. To do so, I examined the percent of bootstrap replicates where treatment effects are positive, zero, or negative by quantile for the reading post-test comparing TFA teachers to all non-TFA teachers, which I show in Figure A.3. This figure examines treatment effect estimates in each of the 999 replicates, which allows us to determine whether there are similar trends in the direction of the effects at different portions of the distribution. The confidence intervals from the estimates presented in the manuscript are determined to be significant at the 5% level (two-tailed test) if 97.5 percent of them are in the same direction. While it is clear that at certain portions of the distribution this threshold is not met, it is also clear when examining the replicates that there are distinctive and countervailing trends at the bottom and the top of the distribution, suggesting that there are different effects at different portions of the distribution.

Comparison with other distributional research on TFA

Concurrent with this paper, which was presented in February 2013 at the annual Sociology of Education Association meeting, in March at the spring meeting for the Society for Research on Educational Effectiveness, and in May at the annual meeting of the American Educational Research Association, a working paper released by Antecol et al. (2013) in April

also examines distributional differences in the effects of TFA on student achievement in elementary school using the MPR data (since published as Antecol, Eren, & Ozbeklik, Forthcoming).

Antecol et al. find evidence of positive effects throughout the distribution of mathematics and no effects at any point along the distribution of reading. Their method differs somewhat from my approach. Their distributional estimation approach follows the work of Canay (2011) and uses fixed effects quantile regression (FEQR) to estimate treatment effects using fixed effects for randomization block. In doing so, they estimate conditional quantile treatment effects. While conditional quantile treatment effects are potentially of interest, they do not necessarily correspond to the unconditional quantile treatment effects. Firpo, Fortin, & Lemieux (2009) note that conditional quantile effects, such as those produced from quantile regression models "cannot be used to answer a question as simple as 'what is the impact on median earnings of increasing everybody's education by one year, holding everything else constant?'" (pg. 1), and show that the effect of union membership on log wages at the conditional 90^{th} percentile is positive, but that the effect at the unconditional 90^{th} percentile is negative.

Intuitively, we can think of this as follows: If we took a population and split it into a number of groups along demographic or other characteristics, the overall mean would correspond to the weighted average of the means of each of the subgroups. However, the same is not true of the 90^{th} percentiles, where the 90^{th} percentile of the overall distribution does not necessarily correspond to either the mean of the subgroups' 90^{th} percentiles or to the 90^{th} percentile of their 90^{th} percentiles. Thus, while conditional quantile effects can net out the effects of other variables, they do so at the expense of rendering the underlying distribution difficult to understand intuitively; it is unclear how to think about the conditional 90^{th} percentile, and as

such, it is difficult to know how to interpret the results. Firpo et al. (2009) offer one solution, providing a re-centering function that allows researchers to translate between the conditional and unconditional quantile effects. As noted above, here I take a different approach, using propensity score weighting to provide estimates that adjust for differences while still providing results that correspond to an underlying distribution that lends itself to an intuitive interpretation (Firpo, 2007).

By using inverse propensity score weights, I can also adjust for potential baseline differences, and not just for the fact that randomization took place within block (school and grade level). The weights used also adjust for treatment group differences in demographic characteristics, the proportion of the classroom that was in the research sample, the presence of missing data, and attrition. Most importantly, the weights adjust for the fact that random assignment did not produce balance between treatment and control across the distribution of either the math or reading at baseline, which fixed effects alone would not do.

It is also worth noting that Antecol et al (2013) include first graders in their analysis sample. I do not include first graders because their reading tests are scored differently from the tests of $2^{nd} – 5^{th}$ graders, though as mentioned, my supplemental analyses indicate that including first graders in the model does not substantively affect my results.

[40] As a result of bootstrapping the 999 samples with replacement, the point estimates are not always centered within symmetric confidence intervals. In portions of the distribution where there is great score density, for example, many students with the same or similar score, it is likely that a given percentile will have well defined, tight confidence intervals, which may even be equal in value to the point estimates. In contrast, in certain portions of the distribution with relatively few students, confidence intervals are more likely to be far from the point estimates. In addition, if there are large score jumps that can fall in different percentiles depending on the particular bootstrap, this can lead to particularly large changes in the confidence intervals from one percentile to the next (as is the case with the reading CI near the 10^{th} percentile). It can also be possible that these score changes happen at margins which lead the point estimates to change, but result in changes in the confidence intervals across several percentiles. This occurs because across the bootstraps, the same highest or lowest scores are consistently drawn as the 2.5^{th} and 97.5^{th} highest scores across several percentiles even though the point estimate from the experimental data fluctuates.

Table A.1

Balance across TFA and non-TFA classrooms on missing values with and without inverse propensity score weights[1]

	Control Mean	T-C Difference	SE	P-value[2]
Differences in prevalence of missing scores unweighted				
Missing Math Pretest	0.022	-0.022	0.006	0.000
Missing Reading Pretest	0.000	0.028	0.006	0.000
Invalid (99) score on Math Pretest	0.027	0.024	0.010	0.020
Invalid (99) score on Reading Pretest	0.057	-0.026	0.011	0.018
Missing Math Posttest	0.001	0.032	0.007	0.000
Missing Reading Posttest	0.000	-	-	-
Invalid (99) score on Math Posttest	0.021	0.001	0.008	0.905
Invalid (99) score on Reading Posttest	0.005	0.001	0.004	0.802
Differences in prevalence of missing scores using inverse propensity score weights				
Missing Math Pretest	0.012	-0.012	0.012	0.313
Missing Reading Pretest	0.000	0.013	0.012	0.304
Invalid (99) score on Math Pretest	0.040	-0.003	0.008	0.714
Invalid (99) score on Reading Pretest	0.048	-0.004	0.013	0.785
Missing Math Posttest	0.002	0.015	0.016	0.358
Missing Reading Posttest	0.000	-	-	-
Invalid (99) score on Math Posttest	0.022	-0.002	0.007	0.775
Invalid (99) score on Reading Posttest	0.006	0.000	0.004	0.994
Sample Size: N = 1430				

[1] Sample includes 2nd - 5th graders with at least one valid pre- and post-test

[2] P-values for weighted comparisons calculated after using inverse propensity score weights and clustering on randomization block

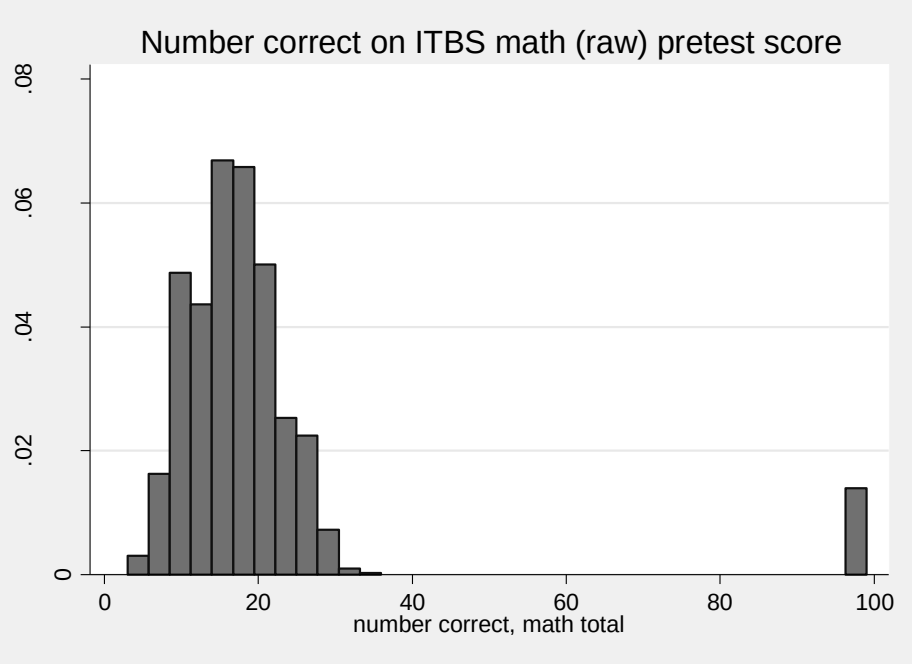
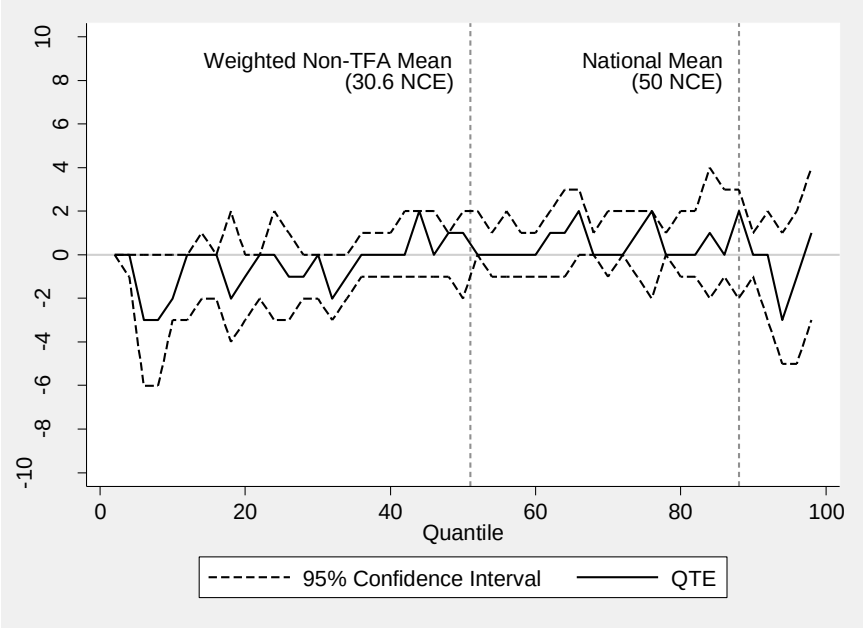


Figure A.1. Histogram of raw ITBS math items correct in fall.

Notes: Figure shows histogram of raw number of math items correct on baseline ITBS as reported in the public-use version of the data. The large point mass at 99 represents those individuals with ITBS raw math scores of 99 and associated National Percentile Ranking scores of 0, and represents a missing data code. Data from the Mathematica Policy Research Evaluation of Teach For America. Includes students in grades 2-5.

Panel A. Difference in Fall Math Achievement



Panel B. Difference in Fall Reading Achievement

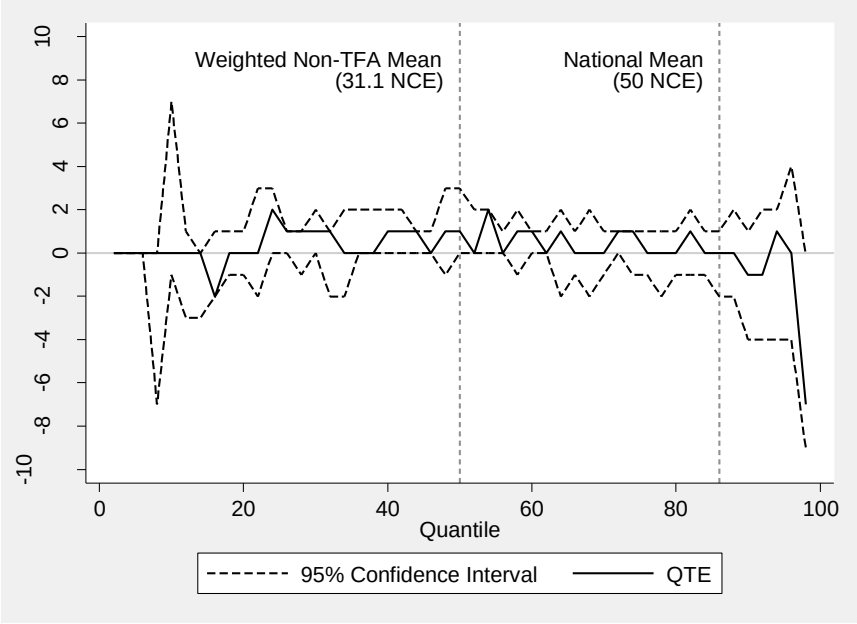


Figure A.2. Inverse propensity score weighted quantile treatment effect estimates of the impact of assignment to TFA classroom on reading Normal Curve Equivalent scores at baseline.

Notes: Panels A & B of figure show QTE for the effect of being assigned to a TFA classroom on math and reading Normal Curve Equivalent scores from the Iowa Test of Basic Skills at baseline. Estimates are weighted using inverse propensity score weights. Weights are $1/\hat{p}$ for treatment observations and $1/(1-\hat{p})$ for control observations, where $\hat{p}$ is generated from a logistic regression of treatment status on baseline demographics, sample design variables, and baseline test score deciles. 95% CIs are obtained by bootstrapping with replacement within randomization block. Data from the Mathematica Policy Research Teach For America Evaluation.

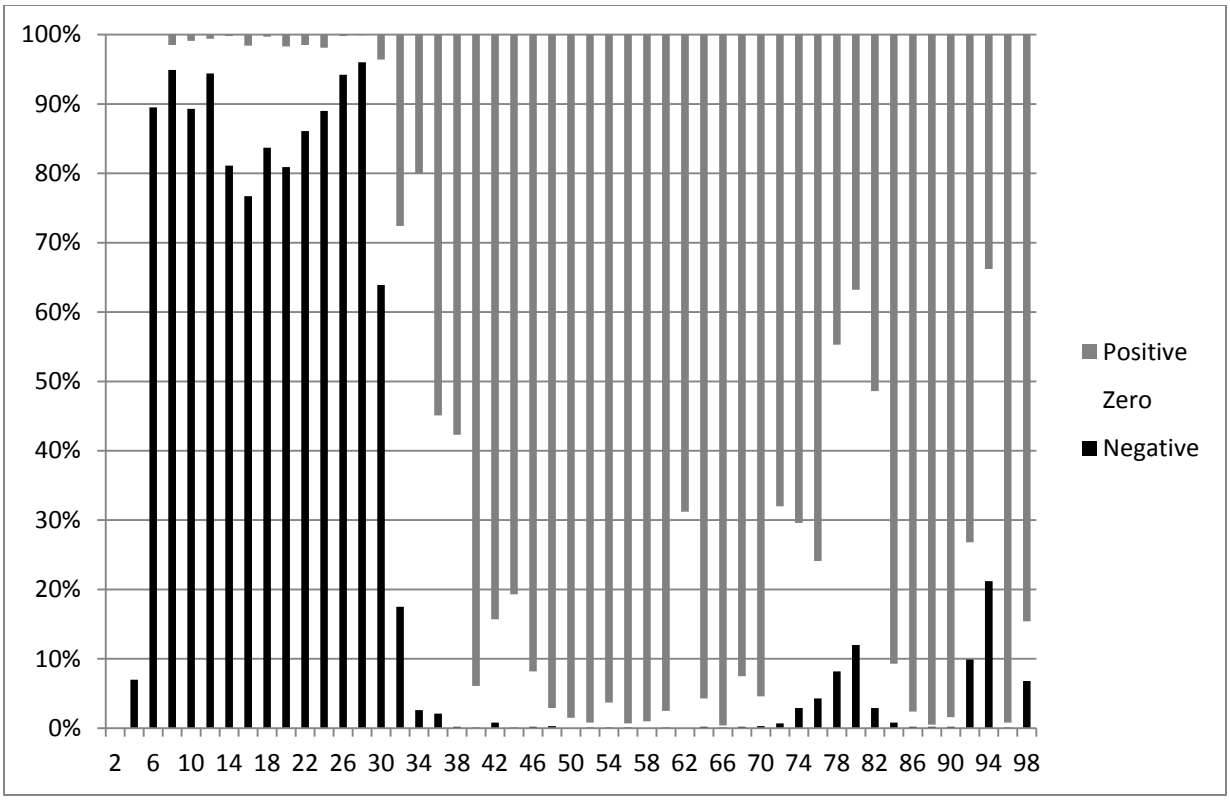


Figure A.3. Percent of bootstrap replicates where treatment effects are positive, zero, or negative by quantile for the reading post-test comparing TFA teachers to all non-TFA teachers.