Teach For America National AmeriCorps Evaluation

2017–18 and 2018–19 School Years

Robert Nathenson, Max Pardo, Sushmita Subedi, Chad Henry, Megha Joshi, Mark Lachowicz, Kyle Neering, Sarah Peko-Spicer, Marina Castro, Adrian Duran, Angela Su, Emily Baumann, and Cassandra Meyer

DECEMBER 2021

American Institutes for Research® | AIR.ORG

AIR®
Advancing Evidence.
Improving Lives.

Teach For America National AmeriCorps Evaluation

2017–18 and 2018–19 School Years

DECEMBER 2021

Authors:

Robert Nathenson, Max Pardo, Sushmita Subedi, Chad Henry, Megha Joshi, Mark Lachowicz, Kyle Neering, Sarah Peko-Spicer, Marina Castro, Adrian Duran, Angela Su, Emily Baumann, and Cassandra Meyer

Acknowledgments

The authors would like to thank the AIR leaders who provided support and thought partnership: Dean Gerdeman, Ellie Fulbeck, Jordan Rickles, and Martyna Citkowicz.

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**

16092_11/21

Contents

Exhibits

Executive Summary

The American Institutes for Research® (AIR®) conducted a quasi-experimental evaluation of student reading and mathematics achievement in eight Teach For America (TFA) regions. Students who had a TFA AmeriCorps Member or alumnus as their primary teacher were compared to a matched sample of students who had a non-TFA teacher. For the analysis, the effect of receiving instruction from a TFA teacher is estimated relative to receiving instruction from a teacher with similar years of experience but who entered the teaching workforce through a non-TFA route (e.g., a university-based training program or alternative teacher preparation program). We examined performance on end-of-grade state assessments controlling for prior-year performance in two recent school years—spring of 2017–18 and spring of 2018–19—in 13 sites within the following eight regions: (1) New York City, NY; (2) Camden, NJ; (3) Charlotte-Piedmont Triad ("CPT"), NC; (4) Eastern North Carolina ("ENC"), NC; (5) an urban district in the Mountain West Region; (6) Clark County, NV; (7) Bay Area, CA; and (8) Los Angeles Area, CA. TFA teachers were matched to non-TFA teachers through a propensity score matching model. The site-level data were then analyzed in a multi-level multivariate regression model with students nested within teachers. In order to synthesize results to estimate an overall effect, we first pooled results from the site-level to the region-level and then pooled results from the eight regions using a meta-analytic approach. Our two primary research questions were as follows:

1. Across TFA regions, what was the impact of being taught by TFA AmeriCorps Members ("ACMs") on student academic achievement (reading and mathematics) as compared to being taught by novice non-TFA teachers?
2. Across TFA regions, what was the impact of being taught by TFA alumni ("Alum") on student academic achievement (reading and mathematics) as compared to being taught by experienced non-TFA teachers?

Key Findings

- **TFA ACMs – No statistically significant average differences for reading or mathematics**
  - Across all regions, being taught by a TFA ACM teacher did not result in significant average differences in student standardized state assessment reading or mathematics scores compared to being taught by a non-TFA novice teacher. This finding holds in both the 2017–18 and 2018–19 school years.
- **TFA Alum – No statistically significant average differences for reading or mathematics**
  - Across all regions, the meta-analysis indicates that being taught by a TFA Alum teacher did not result in significant average differences in student reading or mathematics scores compared to being taught by a non-TFA experienced teacher in either the 2017–18 or 2018–19 school years.

Introduction

For more than 25 years, Teach For America (TFA) has recruited and prepared aspiring leaders to teach in low-income and disadvantaged communities throughout the United States in order to help confront educational inequity through measurable student growth. With more than 55,000 alumni and current members, TFA is one of the largest nationwide nontraditional routes into the teaching profession. Today, TFA continues its commitment of working toward excellence and equity for all by recruiting an increasingly diverse group of teachers.

Although prior research on TFA members has shown that students taught by TFA teachers perform similarly to or better than students in non-TFA classrooms (Antecol et al., 2013; Backes et al., 2016; Clark et al., 2013; Clark et al., 2017; Glazerman et al., 2006; Xu et al., 2011), new research is needed to better understand the impact of the more recent and diverse set of TFA members, as well as the continuous changes TFA has made to its recruitment, selection, and training models. Consequently, TFA has contracted with the American Institutes for Research (AIR) to conduct a rigorous evaluation of recent TFA-AmeriCorps Member cohorts to examine the impact of this more diverse set of members on student learning across multiple regions throughout the country. This report provides a large-scale examination as to how TFA AmeriCorps Members performed based on student reading and mathematics outcomes for the 2017–18 and 2018–19 school years across eight TFA regions—(1) New York City, New York; (2) Camden, New Jersey; (3) Charlotte-Piedmont Triad, North Carolina and (4) Eastern North Carolina; (5) an urban district in the Mountain West Region[1]; (6) Clark County, Nevada; (7) Bay Area, California; and (8) Los Angeles Area, California. For this report, students who had a TFA AmeriCorps Member or alumnus as their primary teacher were compared to a matched sample of students who had a non-TFA teacher. For the analysis, the effect of receiving instruction from a TFA teacher is estimated relative to receiving instruction from a teacher with similar years of experience but who entered the teaching workforce through a non-TFA route (e.g., a university-based training program or alternative teacher preparation program).

Research Questions

This study addressed the following research questions:

[1] The Mountain West Region requested anonymity when reporting results.

1. Across TFA regions, what was the impact of being taught by TFA AmeriCorps Members ("ACMs") on student academic achievement (reading and mathematics) as compared to being taught by novice non-TFA teachers?
2. Across TFA regions, what was the impact of being taught by TFA alumni ("Alum") on student academic achievement (reading and mathematics) as compared to being taught by experienced non-TFA teachers?

Population of Interest

In this study we focus on achievement of students taught by TFA ACMs and Alum as compared to those taught by non-TFA novice and experienced teachers. We focus on students in grades 4-8 as these students are required to take an end of year state assessment in reading and mathematics, and were required to do so in the previous year.[2]

Sources of Data

We used data from two sources for this study. First, lists of TFA ACMs and Alum in the 2017–18 and 2018–19 school years in each of the TFA regions examined were provided by TFA. These lists also included the year a teacher started with TFA (their TFA start year). Second, student- and teacher-level data were largely provided by research departments of districts/charter management organizations (CMOs) within the eight TFA regions – 13 different sites in total.[3]

Data received across districts/CMOs included student end-of-year state assessment outcome data, student demographic information, attendance rates, course enrollment, teacher demographic information, and student-teacher linkages through course enrollment. Staff at each site (i.e. district/CMO/NCERDC) matched the TFA ACMs and Alum lists to their database, flagging in each academic year whether a teacher was a TFA ACM, TFA Alum, or neither.

Methodology

Staff at AIR cleaned and analyzed the data received from each site. For a detailed discussion of the data-cleaning process, including how the primary teacher for each subject was identified, please see Appendix A.

[2] The only exception is Camden, for which high school students are also required to take state assessments in reading and mathematics.

[3] Within the New York, Camden, Mountain West district, and Clark County TFA regions, information was provided by the eponymous district. In the CPT and ENC regions in North Carolina, information for all districts in both regions was provided by Duke University's North Carolina Education Research Data Center (NCERDC). Information for the LA Area and Bay Area came from multiple districts/CMOs. The Los Angeles region included data from Los Angeles Unified School District (LAUSD), KIPP SoCal, and Green Dot Public Schools. In the Bay Area region we received data from Oakland Unified School District, (OUSD) West Contra Costa School District (WCCSD), KIPP Northern California, and Aspire Public Schools.

Outcomes

End-of-Grade State Assessment. For comparability, we standardized student end-of-grade state assessment scores by grade and year in reading and mathematics. In this way, the distributions of student performance are preserved and placed into a common metric. This measure of student performance is continuous.

Defining Novice and Experienced Teachers

As described earlier, the focus of this study is to evaluate the effect of receiving instruction from a TFA teacher relative to receiving instruction from a teacher with similar years of experience but who entered the teaching workforce through a non-TFA route (e.g., a university-based training program or alternative teacher preparation program). We split years of experience into two groups, where novice teachers are defined as teachers with 1-2 years of experience and experienced teachers as those with at least 3 years of experience.

To compare TFA ACMs and TFA Alum to similarly experienced non-TFA teachers, we first had to identify a measure of years of experience. Most sites provided teacher's years of experience in their school or district.[4] For TFA ACMs, we define years of experience by subtracting the year a teacher started with TFA (as provided by TFA) from the academic school year.[5] This provided a time-varying measure of years of experience for TFA ACMs derived from TFAs records of their ACMs' start year. For all non-TFA teachers and TFA Alum, we used each sites' years of experience or salary step to define years of experience (salary step comes from a predetermined annual teacher salary table, where a teacher's step, or level, is usually determined from a combination of years of experience / years of service and education credentials and credits). For further details on how years of experience was measured, please see Appendix A.

Propensity Score Matching

Because this evaluation focuses on the effect of receiving instruction from a TFA teacher relative to receiving instruction from a teacher with similar years of experience but who entered the teaching workforce through a non-TFA route, we match TFA ACMs to novice non-TFA teachers and TFA Alum to experienced non-TFA teachers. We employed a two-stage propensity score matching procedure (Rickles, 2014). This matching was conducted separately

[4] OUSD and WCCSD provided salary step instead. In OUSD, salary step represents the "years of service as a full-time certified teacher" (https://www.ousd.org/cms/lib/CA01001176/Centricity/Domain/105/OUSD%20Teacher%20Salary%20Scale%20Printable.pdf). In WCCSD, teachers "receive yearly step…increases commensurate with their teaching and education experience" (https://www.wccusd.net/cms/lib/CA01001466/Centricity/domain/53/salary%20schedules/2021-22/UTR%208-2021-22%20-8.4.21.pdf).

[5] Experience for TFA ACMs was calculated using the difference between the current school year and the TFA start year, increased by one. For OUSD, which didn't include TFA start year. We used salary step instead.

for reading and mathematics. For each subject, teachers were exact matched within specified grade bands (4-5, 6-8, 9-12) and teacher experience: novice teachers (≤2 years) and experienced teachers, where experienced teachers were separated into two categories of those with 3-5 years of experience and those with at least 6 years of experience. Teachers were matched on their aggregate prior year student achievement, and the proportion of students that were students of color, female, special education, and English Learners (ELs). The procedure used nearest neighbor 1:1 matching with replacement and a caliper of 4.0. In the first stage, TFA teachers were matched to non-TFA teachers within the same school. For TFA teachers that did not match with a teacher in the same school, TFA teachers were matched to non-TFA teachers in different schools within a school cluster. School clusters were defined by comparability of school-average standardized prior year achievement scores. The average match rate was 80%, with a region minimum of 57% and maximum of 100%. This process resulted in 598 matched TFA teachers in reading and 627 in mathematics.

Estimation Method

Post-matching, the impact on student achievement for the group of students who were taught by TFA teachers as compared to the matched sample of students who were taught by non-TFA teachers represents the average treatment effect on the treated. To estimate this impact, we fit a two-level hierarchical linear model with students (Level 1) nested within teachers (Level 2). The model included a vector of individual student characteristics and a vector of teacher-level controls. Student characteristics included grade, race/ethnicity, sex, EL status indicator, special education status indicator, and free or reduced-price lunch eligibility status indicator. It also included an indicator for students who repeated a grade, an indicator for students who were transient (i.e., they were not enrolled for the full school year), and a measure of students' attendance rate (defined as the proportion of enrolled school days in which they attended).[6] At the teacher level, the model included teachers' highest level of education, their number of students, and the proportion of their students who were the following: students of color, female, special education students, ELs, and eligible for free or reduced-price lunch. It also included aggregate prior year student achievement at the school level and school size.[7] We used robust standard errors. The model included weights that account for some TFA teachers being matched to the same comparison teachers (due to matching with replacement).[8] TFA ACMs were compared to non-TFA teachers with 2 or fewer years of experience. TFA Alum were compared to non-TFA teachers with at least 3 years of experience. We ran separate models for each of the 13 sites in each year (2017–18 and 2018–19), for each subject (reading and

[6] We note that not all pieces of information were provided in all sites.
[7] School size was not included in the NYC analyses.
[8] Weights and robust standard errors were not included in the Camden analyses due to small sample size.

mathematics), and for each TFA category (TFA ACMs versus novice non-TFA and TFA Alum versus experienced non-TFA). In total, we estimated 104 regression specifications.

Meta-Analysis

We used two meta-analytic models to synthesize the results from 13 sites to estimate an overall average treatment effect across 8 regions of having a TFA teacher on reading and mathematics student achievement in the 2017–18 and 2018–19 school years. (1) We first pooled results from the site-level to the region level in the LA and Bay Areas. (2) We then pooled results from the eight regions. To pool effects by region, we ran a univariate fixed-effects meta-regression model. To estimate an overall average across all regions, we ran a random-effects multi-level meta-regression model with districts nested within regions (Bloom et al., 2017; Hedges & Olkin, 2014; Hedges & Vevea, 1998).

Who Was Included in the Study?

The analytic sample of TFA teachers included in this study was restricted to those that were linked by reading or mathematics subject to a student, that were identified as the primary teacher in that subject for that student, and had enough students ($\geq$5) to calculate aggregate student information. Non-TFA teachers were similarly restricted, and were only included if they were matched to a TFA teacher through the propensity score matching process described previously. Students were restricted to those without missing achievement score, prior year achievement score, and covariate data, and who were able to link to a primary reading or mathematics teacher.

Exhibits 1a and 1b provide further detail on how we derived the number of teachers included in the analytic sample for the reading and mathematics analyses. As described previously, sites matched the lists of ACMs and Alum provided by TFA to their teacher database. The first column ('Original') details the output of that match - the total number of teachers that were flagged as TFA within each region, by year and ACM/Alum status. For instance, 551 teachers in NYC in both Exhibits 1a and 1b. The next column reflects the number of teachers that were linked to students by subject (e.g. 531 in reading and 400 in mathematics in NYC). The following are those we identified as the student's primary teacher, with the subsequent column teachers that had sufficient data on the variables needed to conduct the matching. The final column represents the number of TFA teachers that were successfully matched to non-TFA teachers. Overall 80% of the TFA teachers eligible for inclusion in the matching were matched to a non-TFA teacher. This final column often represents the set of TFA teachers that were included in the analysis, though some teachers in some regions were excluded post-matching if their

students were missing the information described above. We note that there were few differences between the pre-matched and analytic samples on student, teacher, and school-level characteristics (described further below).

Exhibit 1a. Number of TFA Teachers by Data Processing Stage (Reading)[9]

Region	Original (N=3434)[10]	Linked to Students (N=2190)	Primary Teacher (N=1953)	Sufficient Data (N=758)	Matched (N=598)
NYC, NY	551	531	379	160	121
Camden, NJ	3	3	3	3	3
CPT, NC	491	299	283	120	119
ENC, NC	311	207	192	98	98
Clark County, NV	705	336	323	122	71
Bay Area, CA	768	350	322	165	129
LA Area, CA	605	464	451	90	57

Exhibit 1b. Number of TFA Teachers by Data Processing Stage (Mathematics)[11]

Region	Original (N=3434)[12]	Linked to Students (N=1987)	Primary Teacher (N=1854)	Sufficient Data (N=764)	Matched (N=627)
NYC, NY	551	400	296	119	85
Camden, NJ	3	N/A	N/A	N/A	N/A
CPT, NC	491	324	317	189	189
ENC, NC	311	187	181	91	85
Clark County, NV	705	322	320	127	73
Bay Area, CA	768	319	308	162	144
LA Area, CA	605	435	432	76	51

[9] Data for the Mountain West district was previously processed for a separate TFA report.
[10] Combines all TFA teachers across subjects.
[11] Data for the Mountain West district was previously processed for a separate TFA report.
[12] Combines all TFA teachers across subjects.

Reading. The analytic sample for novice teachers (≤2 years of teaching experience) included 265 ACMs (see Exhibit 2a below) with 10,570 linked students (Exhibit 3a) and the comparison group of 211 non-TFA teachers with 8,072 linked students. The analytic sample for experienced teachers included 328 TFA Alum with 14,098 linked students and the comparison group of 287 non-TFA teachers with 11,239 linked students.

Mathematics. The analytic sample for novice teachers (≤2 years of teaching experience) included 293 ACMs (Exhibit 2b) with 12,214 (Exhibit 3b) linked students and the comparison group of 214 non-TFA teachers with 8,443 linked students. The analytic sample for experienced teachers included 293 TFA Alum with 13,993 linked students and the comparison group of 263 non-TFA teachers with 11,298 linked students.[13]

Exhibit 2a. Teacher Analytic Sample Sizes by School Year, Region, and Group (Reading)

Region	Novice Teachers		Experienced Teachers	
	TFA Teachers (ACMs) (N=265)	Non-TFA Teachers (N=211)	TFA Teachers (Alum) (N=328)	Non-TFA Teachers (N=287)
2017–18				
NYC, NY	12	11	52	49
Camden, NJ	1	1	1	1
CPT, NC	28	25	18	18
ENC, NC	21	16	11	11
Mountain West district	17	17	10	7
Clark County, NV	13	10	16	15
Bay Area, CA	27	16	35	27
LA Area, CA	7	7	20	15
2018–19				
NYC, NY	13	11	44	42
Camden, NJ	N/A[14]	N/A	1	1
CPT, NC	32	26	21	21

[13] Though the analytic sample sizes for reading and mathematics are reported separately by year and subject, a teacher could have appeared in multiple years or subjects. If so, they were counted as part of each sample.
[14] There were no TFA ACMs in reading in Camden in 2018-19.

Region	Novice Teachers		Experienced Teachers	
	TFA Teachers (ACMs) (N=265)	Non-TFA Teachers (N=211)	TFA Teachers (Alum) (N=328)	Non-TFA Teachers (N=287)
ENC, NC	29	22	12	11
Mountain West district	17	16	8	4
Clark County, NV	17	13	25	22
Bay Area, CA	22	12	39	31
LA Area, CA	9	8	15	12

Exhibit 2b. Teacher Analytic Sample Sizes by School Year, Region, and Group (Mathematics)

Region	Novice Teachers		Experienced Teachers	
	TFA Teachers (ACMs) (N=293)	Non-TFA Teachers (N=214)	TFA Teachers (Alum) (N=293)	Non-TFA Teachers (N=263)
2017–18				
NYC, NY	7	5	42	40
Camden, NJ[15]	N/A	N/A	N/A	N/A
CPT, NC	43	21	18	17
ENC, NC	25	19	6	6
Mountain West district	22	22	5	5
Clark County, NV	10	7	24	23
Bay Area, CA	22	16	42	30
LA Area, CA	6	6	13	14
2018–19				
NYC, NY	8	7	28	27
Camden, NJ[16]	N/A	N/A	N/A	N/A

[15] There were no TFA ACMs in mathematics in Camden in 2017–18.
[16] There were no TFA ACMs in mathematics in Camden in 2018–19.

Region	Novice Teachers		Experienced Teachers	
	TFA Teachers (ACMs) (N=293)	Non-TFA Teachers (N=214)	TFA Teachers (Alum) (N=293)	Non-TFA Teachers (N=263)
CPT, NC	52	33	24	25
ENC, NC	20	16	9	9
Mountain West district	25	25	4	4
Clark County, NV	16	14	23	20
Bay Area, CA	29	18	38	27
LA Area, CA	8	5	17	16

Exhibit 3a. Student Analytic Sample Sizes by School Year, Region, and Group (Reading)

Region	Novice Teachers		Experienced Teachers	
	Students of TFA Teachers (ACMs) (N= 10570)	Students of Non-TFA Teachers (N= 8072)	Students of TFA Teachers (Alum) (N= 14098)	Students of Non-TFA Teachers (N= 11239)
2017–18				
NYC, NY	211	337	1651	1414
Camden, NJ	5	20	18	29
CPT, NC	1602	1106	700	650
ENC, NC	1114	788	497	533
Mountain West district	435	534	330	213
Clark County, NV	681	406	510	498
Bay Area, CA	148	360	1259	842
LA Area, CA	1425	681	1703	1073
2018–19				
NYC, NY	383	237	1333	1494
Camden, NJ	N/A[17]	N/A	6	38

[17] No students in the analytic sample as there were no TFA ACMs in reading in Camden in 2018–19.

Region	Novice Teachers		Experienced Teachers	
	Students of TFA Teachers (ACMs) (N= 10570)	Students of Non-TFA Teachers (N= 8072)	Students of TFA Teachers (Alum) (N= 14098)	Students of Non-TFA Teachers (N= 11239)
CPT, NC	913	925	1098	960
ENC, NC	1228	760	697	598
Mountain West district	555	810	274	71
Clark County, NV	566	360	1469	1154
Bay Area, CA	368	402	632	403
LA Area, CA	936	346	1921	1269

Exhibit 3b. Student Analytic Sample Sizes by School Year, Region, and Group (Mathematics)

Region	Novice Teachers		Experienced Teachers	
	Students of TFA Teachers (ACMs) (N= 12214)	Students of Non-TFA Teachers (N= 8443)	Students of TFA Teachers (Alum) (N= 13993)	Students of Non-TFA Teachers (N= 11298)
2017–18				
NYC, NY	174	134	1509	1175
Camden, NJ[18]	N/A	N/A	N/A	N/A
CPT, NC	2060	809	936	920
ENC, NC	1268	706	306	361
Mountain West district	508	682	90	106
Clark County, NV	419	240	1386	1147
Bay Area, CA	65	278	539	464
LA Area, CA	844	856	2183	1186
2018–19				
NYC, NY	178	216	1084	1094

[18] No students in the analytic sample as there were no TFA ACMs in mathematics in Camden in 2017–18.

Region	Novice Teachers		Experienced Teachers	
	Students of TFA Teachers (ACMs) (N= 12214)	Students of Non-TFA Teachers (N= 8443)	Students of TFA Teachers (Alum) (N= 13993)	Students of Non-TFA Teachers (N= 11298)
Camden, NJ[19]	N/A	N/A	N/A	N/A
CPT, NC	2792	1253	1210	1116
ENC, NC	849	642	433	453
Mountain West district	755	1102	49	142
Clark County, NV	738	524	1600	1177
Bay Area, CA	390	229	914	786
LA Area, CA	1174	772	1754	1171

Main Findings

This section includes a discussion of the findings of this report. First, to put in context the settings in which TFA teachers and non-TFA teachers work, we highlight descriptive differences in student characteristics in the pre-matched samples between TFA and non-TFA teachers. We also describe differences in teacher characteristics between the two groups, as well as aggregate school-level differences. We describe how these differences vary, if at all, between the pre-matched and analytic samples. Second, for the matched samples we discuss our main student achievement findings in the 2017–18 and 2018–19 school years across all eight regions.

Descriptive Findings

Student Level

Broadly, there are several differences in the demographic characteristics of students who had a TFA teacher and a non-TFA teacher in both reading and mathematics (see Exhibits 4a and 4b). Students taught by TFA teachers were more likely to be in middle school (Grades 6–8) than elementary school (Grades 3–5) as compared to students taught by non-TFA teachers. They tended to have lower prior year achievement, as measured by standardized state assessment scores. They were more racially/ethnically diverse—for reading, 29% of TFA taught students were African American compared to 22% for non-TFA taught students. Just over half (52%) of

[19] No students in the analytic sample as there were no TFA ACMs in mathematics in Camden in 2018–19.

TFA students were Hispanic compared to 46% for non-TFA teachers. The percentages were similar for mathematics. Approximately 82% of reading and mathematics students of TFA ACMs were eligible for free or reduced-price lunch, 4 percentage points higher than those of non-TFA teachers.

Exhibit 4a. Student Descriptive Characteristics from Pre-Matched Sample (TFA Versus Non-TFA), Reading (N=3,092,221)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Prior Year Achievement	-0.19	0.00	-0.19	1.00
Grade 4	0.06	0.12	-0.16	0.32
Grade 5	0.07	0.12	-0.14	0.32
Grade 6	0.13	0.10	0.11	0.30
Grade 7	0.13	0.10	0.10	0.30
Grade 8	0.14	0.10	0.14	0.30
White	0.10	0.18	-0.19	0.36
Black	0.29	0.22	0.18	0.39
Hispanic	0.52	0.46	0.13	0.47
Asian	0.06	0.11	-0.15	0.29
American Indian	0.01	0.01	-0.01	0.08
Native Hawaiian/Pacific Islander	0.00	0.00	0.00	0.04
Two or more races	0.02	0.03	-0.02	0.14
Race Not specified	0.00	0.00	-0.01	0.03
Female	0.48	0.48	-0.02	0.50
English Learner	0.16	0.16	0.01	0.36
Special education student	0.23	0.17	0.16	0.37
Free or reduced-price lunch eligibility	0.83	0.79	0.08	0.39
Student repeating current grade	0.03	0.04	-0.06	0.19
Transient	0.14	0.13	0.03	0.32
Proportion of school year attended	0.91	0.92	-0.08	0.12

SMD = standardized mean difference; SD = standard deviation

Exhibit 4b. Student Descriptive Characteristics from Pre-Matched Sample (TFA Versus Non-TFA), Mathematics (N=2,522,728)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Prior Year Achievement	-0.19	0.00	-0.20	1.00
Grade 4	0.07	0.11	-0.14	0.31
Grade 5	0.07	0.11	-0.13	0.31
Grade 6	0.16	0.10	0.19	0.31
Grade 7	0.15	0.10	0.17	0.30
Grade 8	0.14	0.10	0.14	0.30
White	0.11	0.20	-0.21	0.38
Black	0.29	0.25	0.08	0.42
Hispanic	0.49	0.39	0.19	0.48
Asian	0.09	0.12	-0.11	0.31
American Indian	0.01	0.01	-0.01	0.09
Native Hawaiian/Pacific Islander	0.00	0.00	-0.01	0.04
Two or more races	0.02	0.03	-0.02	0.15
Race Not specified	0.00	0.00	-0.02	0.04
Female	0.48	0.48	-0.01	0.50
English Learner	0.20	0.16	0.09	0.36
Special education student	0.19	0.18	0.01	0.38
Free or reduced-price lunch eligibility	0.81	0.77	0.07	0.40
Student repeating current grade	0.03	0.04	-0.07	0.19
Transient	0.14	0.13	0.03	0.32
Proportion of school year attended	0.91	0.91	-0.07	0.13

SMD = standardized mean difference; SD = standard deviation

Teacher Level

When averaged across teachers, more than 91% of TFA teachers' students were students of color, which was at least 8 percentage points higher than that of non-TFA teachers' set of students (Exhibits 5a and 5b). For those for which we had data, at least 83% of TFA students were eligible for free or reduced-price lunch, 4-5 percentage points higher than for non-TFA teachers. TFA teachers taught slightly more students than non-TFA teachers on average.

Exhibit 5a. Teacher Descriptive Characteristics from Pre-Matched Sample (TFA Versus Non-TFA), Reading (N=51,640)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Average Prior Year Student Achievement	-0.50	-0.32	-0.23	0.83
Bachelor's degree	0.88	0.82	0.11	0.23
Some postbaccalaureate	0.00	0.00	-0.02	0.02
Master's degree	0.12	0.18	-0.10	0.23
Doctorate	0.00	0.00	-0.01	0.05
Students of color	0.92	0.84	0.38	0.20
Female students	0.42	0.44	-0.12	0.18
Special education students	0.41	0.34	0.17	0.37
EL students	0.22	0.21	0.07	0.26
Students eligible for free or reduced-price lunch	0.86	0.81	0.21	0.21
Number of students	40	38	0	48

SMD = standardized mean difference; SD = standard deviation

Exhibit 5b. Teacher Descriptive Characteristics from Pre-Matched Sample (TFA Versus Non-TFA), Mathematics (N=35,743)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Average Prior Year Student Achievement	-0.38	-0.29	-0.12	0.82
Bachelor's degree	0.97	0.90	0.14	0.13
Some postbaccalaureate	0.00	0.00	-0.02	0.02
Master's degree	0.03	0.09	-0.11	0.13
Doctorate	0.00	0.00	-0.02	0.02
Students of color	0.91	0.81	0.43	0.21
Female students	0.45	0.44	0.08	0.18
Special education students	0.34	0.37	-0.10	0.37

EL students	0.22	0.19	0.15	0.21
Students eligible for free or reduced-price lunch	0.83	0.79	0.20	0.20
Number of students	49	44	0	43

SMD = standardized mean difference; SD = standard deviation

School Level

At the school level (Exhibits 6a and 6b), TFA ACMs taught in schools, on average, that served a higher proportion of students of color than those of non-TFA teachers. A greater share of the student bodies of these schools were also from low-income households, with about four-fifths eligible for free or reduced-price lunch, 5% percentage points higher than that of non-TFA teachers.

Exhibit 6a. School Descriptive Demographics from Pre-Matched Sample (TFA Versus Non-TFA), Reading (N=5,030)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Students of color	0.89	0.82	0.36	0.18
Female students	0.49	0.48	0.10	0.07
Special education students	0.18	0.19	-0.04	0.12
EL students	0.17	0.16	0.10	0.14
Students eligible for free or reduced-price lunch	0.81	0.76	0.28	0.18
Number of students	722	720	0	546

SMD = standardized mean difference; SD = standard deviation

Exhibit 6b. School Descriptive Demographics from Pre-Matched Sample (TFA Versus Non-TFA), Mathematics (N=4,582)

Covariate	TFA	Non-TFA	SMD	SD
Students of color	0.87	0.80	0.36	0.19
Female students	0.48	0.48	-0.02	0.07
Special education students	0.18	0.19	-0.05	0.11
EL students	0.18	0.16	0.12	0.14
Students eligible for free or reduced-price lunch	0.77	0.72	0.26	0.18
Number of students	730	734	0	554

SMD = standardized mean difference; SD = standard deviation

We also compared the pre-matched (Tables 4a-6b above) and analytic samples (Appendix Exhibits D3a-D5b) to examine whether there were any differences between the two samples. At the student-level, the difference in average prior year achievement between TFA and non-TFA teachers in the pre-matched sample is eliminated in the reading analytic sample and greatly reduced in the mathematics analytic sample. In addition, a smaller share of students with a TFA teacher was designated as special education in the analytic sample. In mathematics, the share of EL students with a TFA teacher increased slightly. All other student characteristics were within 3 percentage points between the TFA pre-matched and analytic samples. At the teacher level, the gap in teachers' average prior year achievement of their students between TFA and non-TFA teachers is greatly reduced between the pre-matched and analytic samples. This, placed with the similar reduction at the student-level, indicates that the matching did a good job of creating a sample of TFA and non-TFA teachers whose students had similar prior year achievement. For further detail on prior year achievement in the pre- and post-matched samples, please see Appendix Exhibits D1a and D1b. Additionally, the average proportion of TFA teachers' students that were female increased to 48% in the analytic sample, closely paralleling the percentage for non-TFA teachers. The proportion of TFA teachers with a Master's degree decreased in the reading analytic sample and increased in the mathematics analytic sample as compared to the pre-matched samples. At the school-level, there were fewer students, on average, at schools in the analytic as compared to the pre-matched samples. There was also a slightly higher share of students designated as EL. All other school characteristics were within 3 percentage points between the pre-matched and analytic samples.

National Test Performance Findings

TFA ACMs – No statistically significant average differences for reading or mathematics

Exhibit 7a details the estimated average treatment effect of standardized reading and mathematics achievement for the group of students taught by TFA ACMs as compared to those taught by the matched sample of novice non-TFA teachers. The last row of the exhibit details the eight region aggregated average treatment effect from the meta-analytic models. All other rows represent the region-specific treatment effect. Dots represent point estimates and lines represent 95% confidence intervals. Colors indicate whether results are from the 2017–18 school year (green) or the 2018–19 school year (orange).

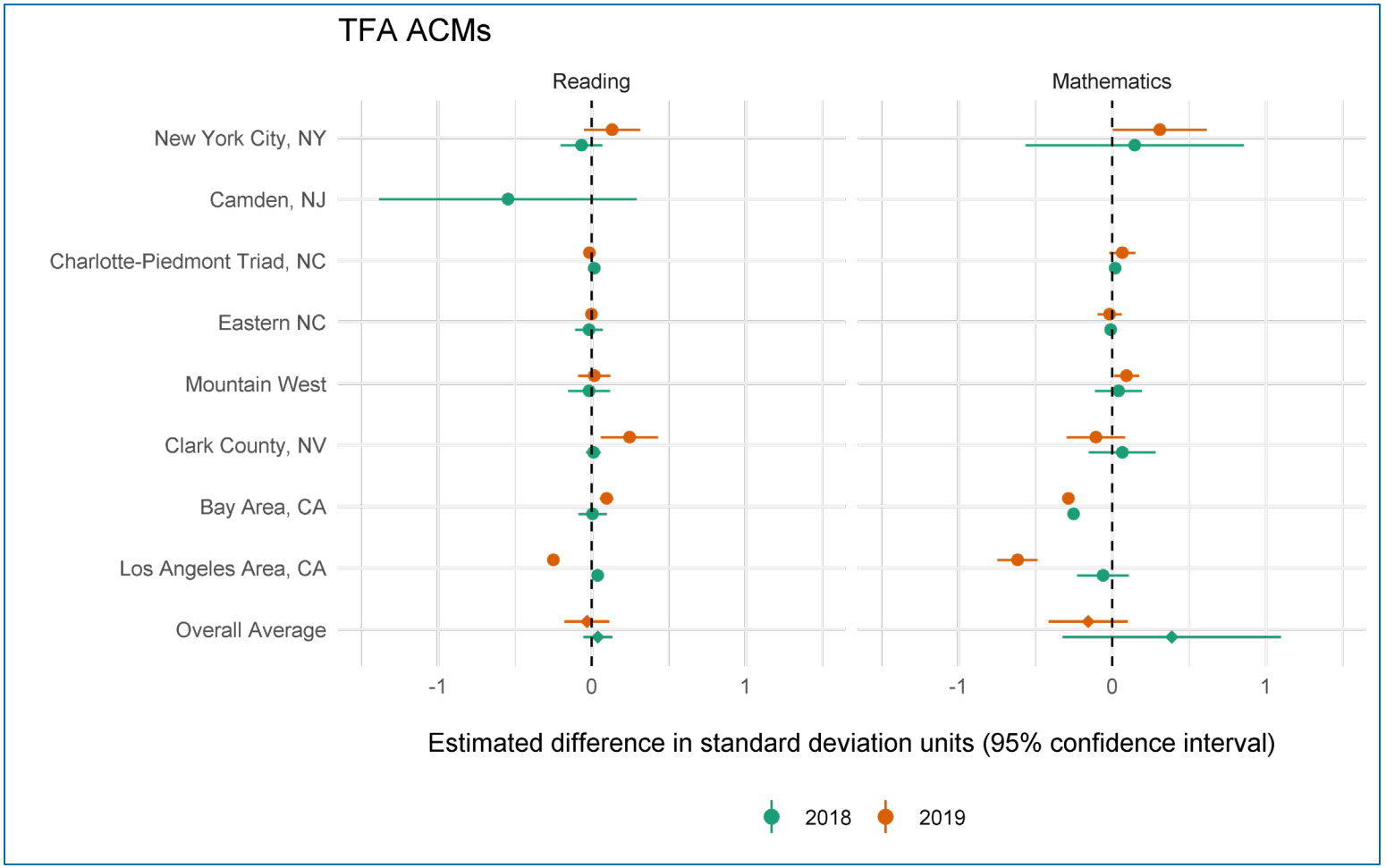
Overall, the meta-analytic average treatment effects for reading and mathematics in the 2017–18 and 2018–19 school years were not statistically significant for students taught by TFA ACMs. While there is some variation across regions, the overall effect indicates that, on average, being taught by TFA ACMs did not result in a statistically significant difference in standardized achievement as compared to being taught by non-TFA novice teachers.

TFA Alum – No statistically significant average differences for reading or mathematics

Exhibit 7b details the average treatment effect in reading and mathematics for the group of students who were taught by TFA Alum teachers as compared to those that were taught by the matched sample of experienced non-TFA teachers. The layout of the results by region and overall is the same as in Exhibit 7a.

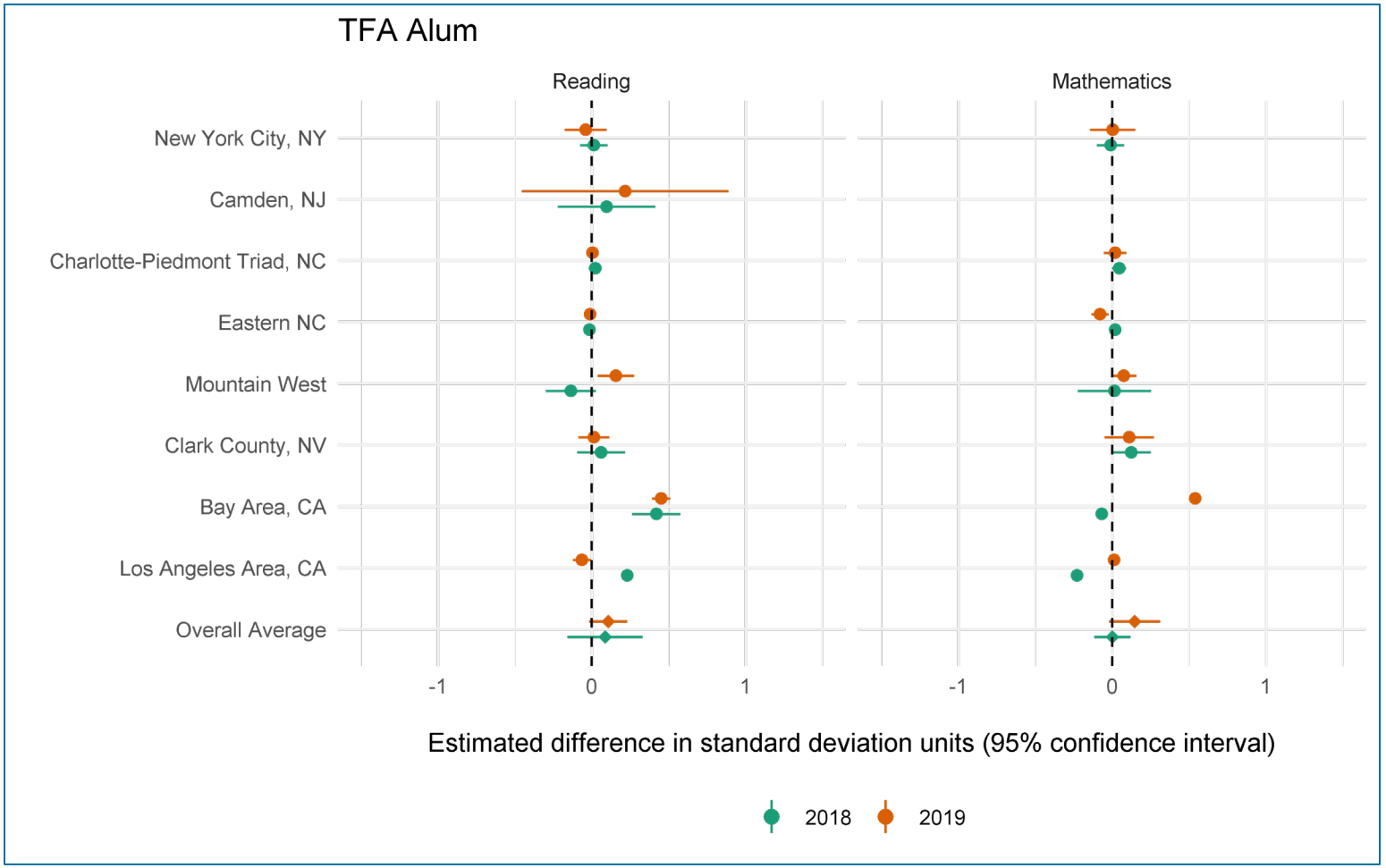
The results for TFA Alum are similar to the findings for TFA ACMs. The meta-analytic average treatment effect for TFA Alum is not significantly different in either year in either subject. That is, there is no statistically significant difference in the overall average treatment effect for reading or mathematics in either year for the group of students who were taught by TFA Alum as compared to the group taught by non-TFA experienced teachers. For further detail on these estimates, please see Appendix D2a (reading) and D2b (mathematics).

Exhibit 7a. Meta-Analysis Results for TFA AmeriCorps Members/Novice Teachers, by Subject and School Year +



+ In the analytic sample in Camden, New Jersey, there were no TFA ACMs in reading in 2018–19 or mathematics in either year.

Exhibit 7b. Meta-Analysis Results for TFA Alum/Experienced Teachers, by Subject and School Year [+]



[+] In the analytic sample in Camden, New Jersey, there were no TFA Alum teachers in mathematics in either year.

Study Limitations

The results presented in this report have several limitations. First, propensity score matching, which matched TFA to non-TFA teachers on observed information only, cannot account for the potential of teacher self-selection into TFA or other teacher preparation programs. In addition, we were often limited by small sample sizes with respect to the amount of information that could be included in the matching. Second, in some instances, the propensity score matching resulted in standardized mean differences between TFA and non-TFA teachers that were greater than the What Works Clearinghouse guideline of .25. For further detail on how the matching reduced differences on teachers' aggregated student prior year achievement, please see Appendix Exhibits D1a and D1b. Third, the analytic sample was restricted to a subset of the overall TFA population based on a number of factors described previously. As discussed earlier, we observed a few differences between the pre-matched and analytic samples, including school size, share of students that had a special education designation, share of students designated as EL, and proportion of teachers with a Master's degree.

Fourth, the way years of experience was operationalized varied across sites. Most sites provided a teacher's years of experience in the school or district, while two sites provided salary step instead. In each year, there were several instances of the TFA ACM/Alum designation not aligning with the years of experience identified. As such, TFA ACMs' years of experience were top coded at 2 and TFA Alum's years of experience bottom-coded at 3 to reflect the TFA ACMs and Alum definitions. Additionally, the TFA ACM measure was typically calculated by subtracting the year the teacher started with TFA (provided by TFA) from the school year. We consider this calculation as the most accurate measure available for years of experience for TFA ACMs. However, TFA Alum and non-TFA years of experience was operationalized by information provided by each site. The limitation of having imperfect years of experience data is that, despite these efforts, in some instances we may not be comparing TFA ACMs to novice non-TFA and TFA Alum to experienced non-TFA teachers.

Conclusion

In this study, we examined reading and mathematics achievement of Grades 4–8 students in the 2017–18 and 2018–19 school years in eight regions across the United States who had a TFA ACM or Alumnus as their primary teacher as compared to those who had a non-TFA teacher.

In these settings, the descriptive analysis shows how in recent years TFA ACMs and Alum taught a greater proportion of students of color, students from low-income households, and students who had lower prior year achievement scores. TFA ACMs were located in schools that, as a whole, served more students from low-income households and students of color than non-TFA teachers.

For the matched sample, findings suggest no overall differences in student achievement across TFA and non-TFA teachers, as measured by the state standardized assessment score. Although there is some variation across regions, there is no discernible pattern to these findings across TFA ACMs or Alum, region, or year.

While these results do not suggest TFA teachers in recent years have outperformed non-TFA teachers, on average, on students' standardized state assessments, the results do support the notion that school districts can expand and diversify their teacher workforce through TFA without significantly affecting student performance.

References

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. *Economics of Education Review, 37*, 113–125.

Backes, B., Hansen, M., Xu, Z., & Brady, V. (2016). Examining spillover effects from Teach For America Corps members in Miami-Dade County public schools. *Journal of Teacher Education, 70*(5). https://doi.org/10.1177%2F0022487117752309

Bloom, H. S., Raudenbush, S. W., Weiss, M. J., & Porter, K. (2017). Using multisite experiments to study cross-site variation in treatment effects: A hybrid approach with fixed intercepts and a random treatment coefficient. *Journal of Research on Educational Effectiveness*, *10*(4), 817-842.

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* (NCEE 2013-4015). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Clark, M. A., Isenberg, E., Liu, A. Y., Makowsky, L., & Zukiewicz, M. (2017). *Impacts of the Teach For America Investing in Innovation scale-up.* Mathematica Policy Research.

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: The Journal of the Association for Public Policy Analysis and Management*, *25*(1), 75–96.

Hedges, L. V., & Olkin, I. (2014). *Statistical methods for meta-analysis*. Academic press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological methods*, 3(4), 486.

Konstantopoulos, S., Hedges, L. V., Cooper, H., & Valentine, J. C. (2019). Statistically analyzing effect sizes: Fixed-and random-effects models. *The handbook of research synthesis and meta-analysis*, 245-280.

Rickles JH, Seltzer M. (2014) A Two-Stage Propensity Score Matching Strategy for Treatment Effect Estimation in a Multisite Observational Study. *Journal of Educational and Behavioral Statistics,* 39(6), 612-636. doi:10.3102/1076998614559748

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.

Appendix A. Data Cleaning

We describe our data-cleaning procedure in this appendix.[20] Students were assigned a single race/ethnicity and sex. In instances in which a student's race/ethnicity or sex designation in the data changed across years, the designation in the most recent year available was used. Student eligibility for free or reduced-price lunch, special education status, and EL status were determined by whether the data indicated that the student ever received those services. Students were also assigned a single grade level per year. Students with multiple grade levels in a single year were assigned to the most recently enrolled grade. When the grade selected for each student matched the grade selected the previous year, the student was designated as a grade repeater.

Achievement outcomes were limited to a single score per subject, test, and school year. For instances in which students had multiple scores for the same outcome, the highest score was selected. Students who were enrolled in the district for fewer days than the number of days offered (the median value of all students by school and year) were considered transient. Attendance rate was calculated using the number of days attended divided by the number of days enrolled.

Students were assigned one (primary) teacher for each subject. In cases in which students attended multiple courses in one subject, a series of rules determined which teacher would be selected as the primary teacher. Each of the following rules were implemented until a single teacher was selected:

1. Course withdrawal date was used to determine the most recent course attended by the student for the school year.
2. Course entry and withdrawal date were used to calculate which course the student attended the longest.
3. The number of students who took the course was used to determine which course was the most common.
4. As a final tiebreaker, we used the first course in the data file, assigned before any data manipulation.

[20] Data for the Mountain Region was cleaned for a different TFA project and only reanalyzed to provide separate estimates for the 2017-18 and 2018-19 years.

For instances in which multiple teachers were assigned the same course for the same duration for the same student, the teacher connected with more students overall and with a greater number of years of experience was selected.

Teacher race/ethnicity and sex followed the same rules as the student variables, with only a single designation across all years. In cases in which teacher education level was missing, the teacher was coded as having a bachelor's degree. Teachers were assigned a single school per subject. In cases in which a teacher was the primary teacher for students across multiple schools, the most common school was selected.

In addition to the process described earlier for measuring years of experience, we conducted the following data cleaning steps for years of experience. We ensured that years of experience was time-varying and that TFA status was time-varying, with TFA teachers switching from TFA ACM to TFA Alum status after year 2. We coded teachers in their first year as having 1 year of experience, in their 2nd year as 2 years, and so on.[21] We do so as we measured student achievement in the spring of each academic year, when a teacher was nearing completion of the school year. We rounded all years of experience to whole numbers when necessary (e.g. NYC provided partial years of experience). For all TFA ACMs in a given year, we top coded their years of experience at 2. For all TFA Alum in a given year, we bottom coded their years of experience at 3.

To validate the operationalization of TFA ACMs experience, across a variety of sites we correlated this measure as compared to the years of experience measure and teachers' age provided by sites. The correlation was usually high. For instance, the correlation between district years of experience and TFA ACMs operationalized years of experience was .98 in Green Dot, .95 in CCSD, .95 in LAUSD, over .99 in Camden, and over .99 in NYC.[22] It was lower in WCCSD at .53, which provided salary step instead. In some sites we examined teacher age in addition to years of experience – the correlation was .86 between age and operationalized TFA ACMs years of experience in Aspire. The correlations were also lower in North Carolina and so in both North Carolina regions we included an additional restriction when operationalizing TFA ACMs and Alum years of experience.[23]

[21] Therefore, sites that assigned zero years of experience to new teachers had the years of experience increased by one for all teachers.

[22] After removing three outlier cases in NYC.

[23] To validate years of experience in North Carolina, we compared the provided measure for TFA ACM and Alum with our calculated years of experience measure. For TFA ACMs, we operationalized years of experience by subtracting the year a teacher started with TFA (as provided by TFA) from the academic school year. For TFA Alum and non-TFA teachers, we calculated years of experience by subtracting the year a teacher graduated with their highest degree (provided in the NCERDC data) from the academic school year. The correlation between North

Indicators for ACM and Alum TFA status were mutually exclusive by year. Additionally, teachers could never be indicated as non-TFA regardless of year if they were ever flagged as such. In cases in which the data provided did not meet these requirements, the TFA indicators were recoded using TFA start year.

Carolina's and our years of experience measures for TFA ranged between .64 and .69 for TFA ACMs (reading and mathematics, respectively), .39 and .64 for TFA Alum, and .73 for non-TFA. Because the measures were not highly correlated (often defined as a correlation of .8 or above), we imposed an additional restriction for defining years of experience. To be included in the analytic sample, novice teachers (defined as those with 2 or fewer years of experience) had to be listed as 0–2 years of experience on both measures. Experienced teachers were those with at least 3 years of experience.

We separated experienced teachers into two categories—those with 3–5 years of experience and those with at least 6 years of experience. For both the 3–5 and 6+ categories of experienced teachers to be included in the analytic sample, we imposed the same restriction—teachers had to fall within the appropriate category under both measures. While this restriction reduced the number of teachers included in the analyses, given the lower correlation between the 2 years of experience measures, it increased the validity of how the years of experience measure was defined. This factor is of critical importance in a study of this nature, which seeks to compare student performance across categories of teachers with very different paths into the teaching profession.

Appendix B. Propensity Score Matching Procedure

We employed a two-stage propensity matching procedure (Rickles, 2014). This matching was conducted separately for reading and mathematics. For each subject, teachers were exact matched within specified grade bands (4-5, 6-8, 9-12) and teacher experience: novice teachers (≤2 years) and experienced teachers, where experienced teachers were separated into two categories of those with 3-5 years of experience and those with at least 6 years of experience. Teachers were matched on the teacher-level variables described previously, including aggregate prior year student achievement, and the proportion of students that were students of color, female, special education, and EL. The procedure used nearest neighbor 1:1 matching with replacement and a caliper of 4.0. In the first stage, TFA teachers were matched to non-TFA teachers within the same school. For TFA teachers that did not match within their school, TFA teachers were matched to non-TFA teachers in different schools within a school cluster. School clusters were defined by comparability of school-average standardized prior year test scores. The total number of school clusters was determined by the number of schools in a district, where districts with less than 12 schools had a single school cluster, districts with 13 to 19 schools had 3 school clusters, and districts with 20 or more schools had 5 school clusters. Cutoff values to determine cluster membership were terciles of the school-average prior year test score distribution for districts with 3 clusters, and quintiles for districts with 5 clusters (there were no cluster cutoffs for districts with a single cluster). The average match rate was 80%, with a region minimum of 57% and maximum of 100%.

Propensity scores were estimated as follows. The estimation procedure was conducted separately for teacher subject (math and ELA) and school year (2018 and 2019). Under the logit function, *g*, let p_{ij} represent the probability teacher *i* is TFA in school *j* and $log\left(\frac{p_{ij}}{1-p_{ij}}\right)$ the log odds. The probability of each outcome can be specified as a linear combination of the log odds of the logit function. Under the below specification, β_1 represents the parameter for a vector of teacher-average variables (*X*), including percentage of minority students (Black, Asian, Latinx, other), percentage of female students, percentage of special education students, percentage of EL students, and average prior year standardized test score. School random effects u_{0j} were included to account for potential dependency of propensities of students within schools, where $u_{0j} \sim N(0, \tau^2)$.

$$g(p_{ij}) = logit(p_{ij}) = log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_1(X)_{iJ} + u_{0j} + \varepsilon_{ij} \quad (1)$$

We used Bayesian generalized mixed models to estimate propensity scores because the specification of prior distributions resulted in more stable estimation, particularly in regions

with relatively few schools (<10). Weak prior distributions were specified for the intercept β_0, regression coefficients β_1, and the between-school random effect variance τ^2. The matching procedure was conducted using the R package MatchIt (version 3.0.2).

Appendix C. Model Specification

Region Analyses

To estimate the effect of having a TFA teacher, we separately analyzed 108 samples: two school years (2017–18 and 2018–19) for two TFA categories (TFA ACMs versus novice non-TFA and TFA Alum versus experienced non-TFA) for two subjects (reading and mathematics) at 13 sites across 8 regions. For each sample of matched TFA and non-TFA teachers, we specified a two-level (students nested within teachers) hierarchical mixed-effects model with random intercepts at the teacher level. At Level 1, y_{ij} represented the outcome of interest for state standardized assessment score. The key parameter of interest was the TFA indicator (0 = not TFA, 1 = TFA). The analyses of novice teachers were limited to those with 0–2 years of experience, and the comparison was between novice non-TFA teachers (0) and TFA ACMs (1). In the analyses of experienced teachers, the comparison was between experienced non-TFA teachers (0) and TFA Alum (1). It was limited to teachers with 3+ years of experience. The analyses included X_i, a vector of student characteristics, including grade indicators; race/ethnicity (binary indicators for Black, Asian, Hispanic, American Indian/Alaskan Native, Native Hawaiian or Other Pacific Islander, and Two or More Race/Ethnicities, with White students as the omitted category); sex (0 male, 1 female); a continuous attendance rate measure ([0,1]); and binary indicators for special education, English Learner, eligibility for free- or reduced-price lunch, repeating a grade, and whether the student was enrolled in the school for only part of the school year. T_{ij} is a vector of teacher-level characteristics, including aggregated prior year achievement of the teacher's students, teacher's highest level of education (binary indicators for some post-college education, master's, and doctorate/JD, with bachelor's degree the omitted category), and a set of continuous measures aggregated from the student level, including proportion of a teacher's students that were students of color, female, special education students, English Learners, and eligible for free or reduced-price lunch. It also included a count of the number of students each teacher taught. S_{ij} is a vector of school-level information aggregated from all students in the school, including the school-level aggregate student prior year achievement and a count of the number of students in the school. We used robust standard errors. The residual was partitioned into the student-level component, r_{ij} and the teacher-level component, u_{0j}, which were both assumed to be multivariate normal. M1 below specifies the modeling approach.

$$y_{ij} = \beta_0 + \beta_1 X_i + r_{ij} \qquad \text{(M1 Level 1- Student Level)}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} TFA_{1j} + \gamma_{02} T_{2j} + \gamma_{03} S_{3j} + u_{0j} \qquad \text{(M1 Level 2 – Teacher Level)}$$

$$\beta_{1j} = \gamma_{10}$$

Meta Analysis

We conducted propensity score matching within districts and estimated average treatment effects separately in each of the districts. Essentially, we emulated a multisite randomized experimental design and estimated site-specific average treatment effect estimates. We then used meta-analytic models to synthesize the results from the 13 sites to estimate an overall average treatment effect.

To pool effects by region, we ran a univariate fixed-effects meta-regression model. Below let r denote region r, i denote site i, where $i = 1 \ldots, m_r$ total sites in region r, and $\widehat{\beta_{ir}}$ denotes the estimated average treatment effect from site i in region r. For each region the effects were pooled as follows (Konstantopoulos & Hedges, 2019).

$$\widehat{\mu_r} = \frac{\sum_{i=1}^{m_r} w_{ir}\, \widehat{\beta_{ir}}}{\sum_{i=1}^{m_r} w_{ir}}$$

Here, $\widehat{\mu_r}$ is the estimated average treatment effect for region r. The fixed effects weights in meta-analysis are as follows. Below let σ_{ir}^2 denote the variance associated with the estimated average effect in site i in region r.

$$w_{ir} = \frac{1}{\widehat{\sigma_{ir}^2}}$$

To estimate an overall average across all regions, we ran a random-effects multi-level meta-regression model with sites nested within regions (Bloom et al., 2017; Hedges & Olkin, 2014; Hedges & Vevea, 1998). Below, let μ denote the overall average effect. The multi-level random effects meta-analytic model is as follows:

$$\widehat{\beta_{ir}} = \mu + u_r + v_{ir} + e_{ir}$$

Here, $u_j \sim N(0, \tau^2)$ is the deviation between each effect parameter and the population average, $v_{ir} \sim N(0, \omega^2)$ is the deviation between each effect parameter and the region-level average effect, and $e_{ir} \sim N(0, \sigma_{ir}^2)$ is the sampling error for effect i in region r. The random-effects weights account for between-study variance (τ^2), within-study variance (ω^2), and sampling error (σ_{ir}^2).

Appendix D. Additional Data Details

The plots below show standardized mean differences (SMDs) on students' prior year achievement scores aggregated to the teacher level. The red dots represent the SMDs calculated on the sample before matching TFA to non-TFA teachers, and the blue dots represent the SMDs calculated on the sample post-matching. The y-axis contains different regions and the x-axis contains the SMDs. The dashed vertical line on x=0 represents absolute balance. For both reading and mathematics, across all regions, the SMDs are closer to 0 post-matching than pre-matching, indicating that matching resulted in better balance on the teacher-level prior year achievement scores.

Exhibit Appendix D1a. Teacher-Level Standardized Mean Differences Pre- and Post-Matching, Reading

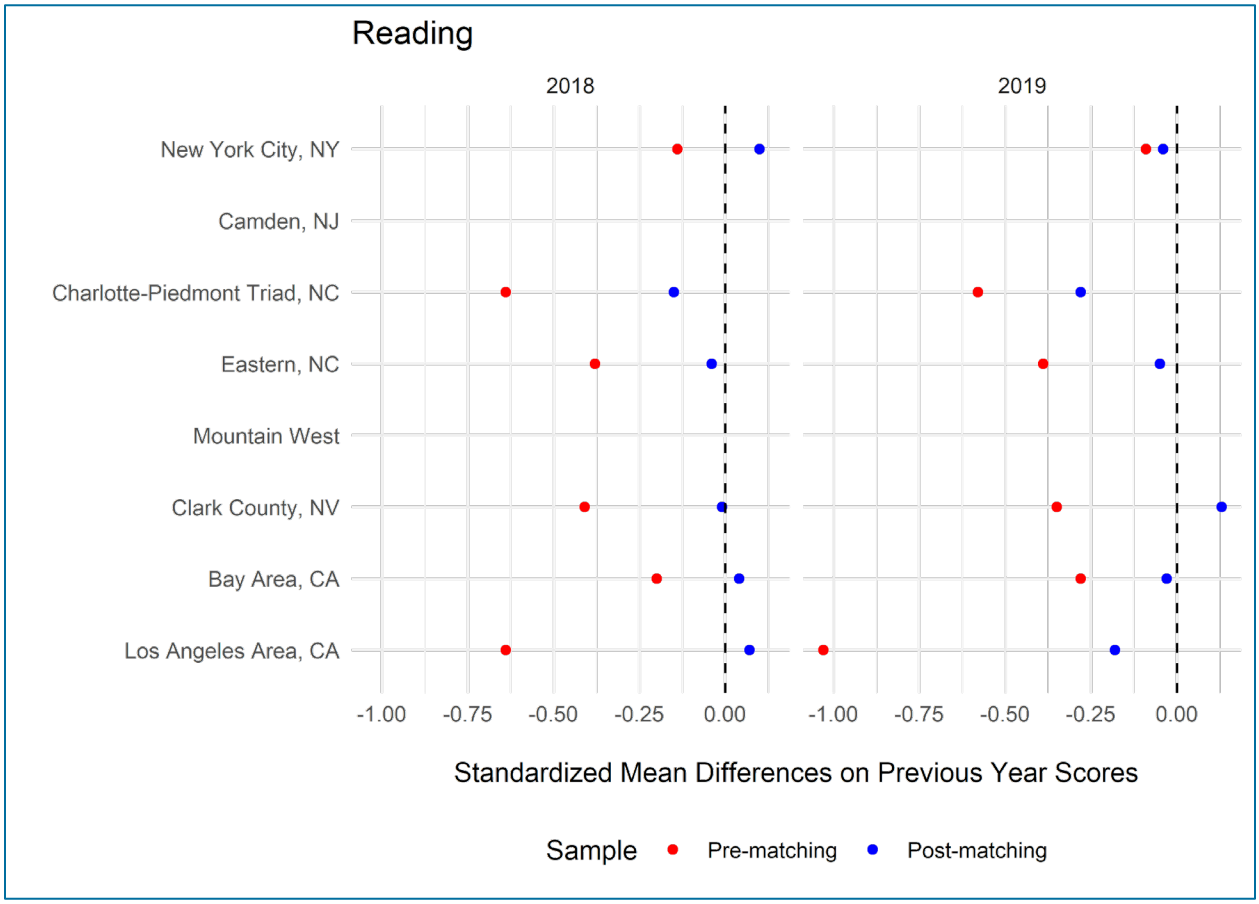


Exhibit Appendix D1b. Teacher-Level Standardized Mean Differences Pre- and Post-Matching, Mathematics

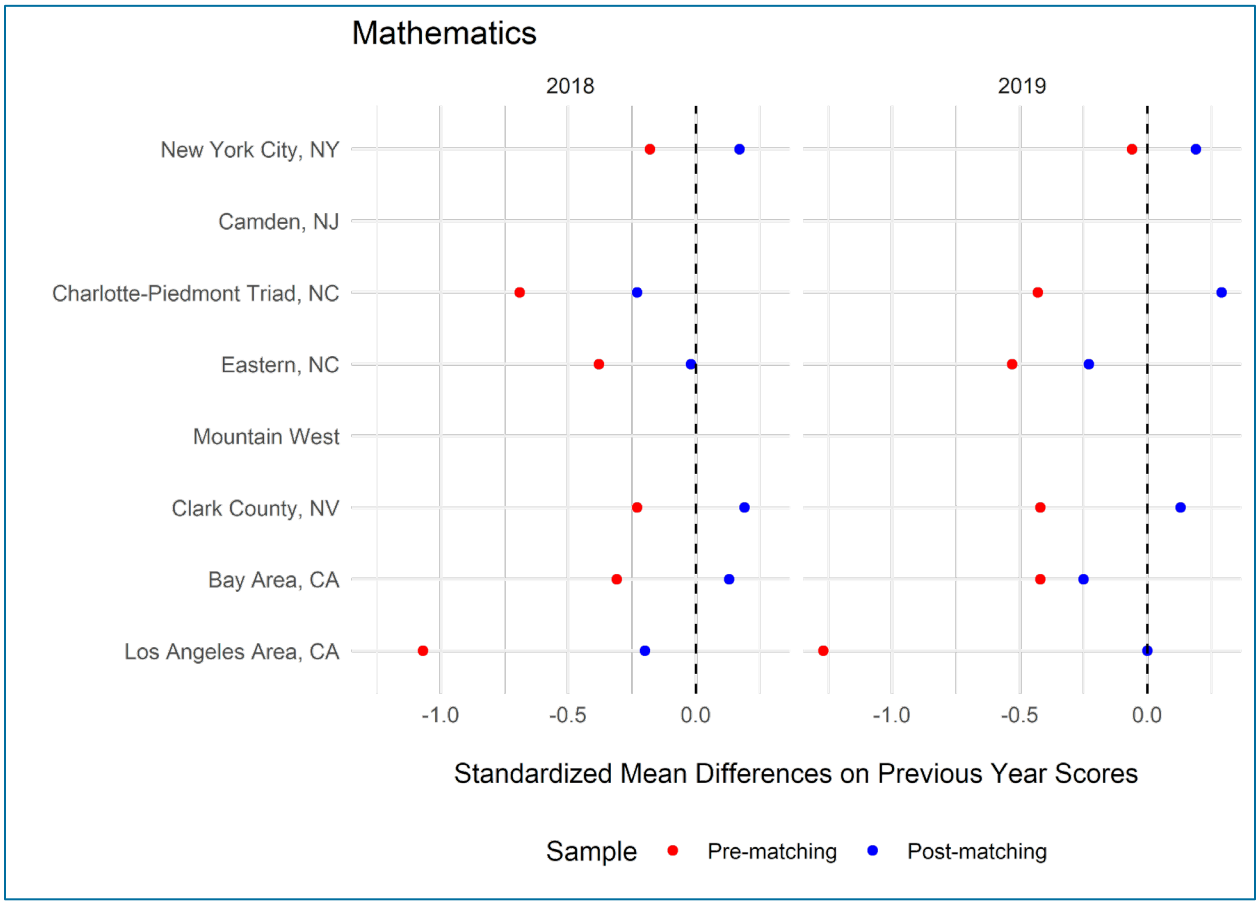


Exhibit D2a. Point Estimates (and Standard Errors) for TFA AmeriCorps Members Analysis by Region, Year, and TFA Status

Subject	2017–18		2018–19	
	Reading	Mathematics	Reading	Mathematics
Reading Overall	.04 (.05)	.39 (.36)	-.03 (.07)	-.15 (.13)
NYC, NY	-.07 (.07)	.15 (.36)	.13 (.09)	.31 (.16)
Camden, NJ	-.54 (.43)	N/A	N/A	N/A
CPT, NC	.02 (.02)	.02 (.02)	-.01 (.01)	.07 (.04)
ENC, NC	-.02 (.05)	-.01 (.02)	.00 (.01)	-.02 (.04)
Mountain West district	-.02 (.07)	.04 (.08)	.02 (.05)	.09 (.04)
Clark County, NV	.01 (.02)	.07 (.11)	.25 (.10)	-.11 (.10)
Bay Area, CA	.01 (.05)	-.25 (.00)	.10 (.02)	-.28 (.02)
LA Area, CA	.04 (.01)	-.06 (.09)	-.25 (.01)	-.61 (.07)

Exhibit D2b. Point Estimates and Standard Errors for TFA Alum Analysis by Region, Year, and TFA Status

Subject	2017–18		2018–19	
	Reading	Mathematics	Reading	Mathematics
Mathematics Overall	.09 (.13)	.003 (.07)	.11 (.06)	.15 (.08)
NYC, NY	.01 (.05)	-.01 (.04)	-.04 (.07)	.00 (.07)
Camden, NJ	.10 (.16)	N/A	.22 (.34)	N/A
CPT, NC	.02 (.02)	.05 (.02)	.01 (.01)	.02 (.04)
ENC, NC	-.02 (.01)	.02 (.01)	-.01 (.01)	-.08 (.03)
Mountain West district	-.13 (.08)	.02 (.12)	.16 (.06)	.08 (.04)
Clark County, NV	.06 (.08)	.13 (.06)	.01 (.05)	.11 (.08)
Bay Area, CA	.42 (.08)	-.07 (.01)	.45 (.03)	.54 (.02)
LA Area, CA	.23 (.01)	-.23 (.01)	-.06 (.03)	.01 (.01)

Exhibit D3a. Student Descriptive Characteristics from Analytic Sample (TFA Versus Non-TFA), Reading (N=40,757)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Prior Year Achievement	-0.14	-0.14	0.00	0.97
Grade 4	0.11	0.14	-0.08	0.30
Grade 5	0.13	0.14	-0.02	0.32
Grade 6	0.21	0.24	-0.10	0.40
Grade 7	0.27	0.27	0.00	0.43
Grade 8	0.28	0.21	0.17	0.43
White	0.10	0.11	0.00	0.28
Black	0.30	0.27	0.06	0.40
Hispanic	0.49	0.52	-0.07	0.44
Asian	0.07	0.06	0.01	0.21
American Indian	0.00	0.00	-0.01	0.06
Native Hawaiian/Pacific Islander	0.00	0.00	0.00	0.05
Two or more races	0.02	0.02	0.01	0.14
Race Not specified	0.00	0.00	0.00	0.02
Female	0.49	0.49	0.01	0.50
EL	0.18	0.20	-0.05	0.35
Special education student	0.15	0.15	-0.01	0.35
Free or reduced-price lunch eligibility	0.81	0.81	-0.01	0.36
Student repeating current grade	0.01	0.01	0.01	0.07
Transient	0.13	0.12	0.00	0.25
Proportion of school year attended	0.94	0.94	-0.01	0.06

SMD = standardized mean difference; SD = standard deviation

Exhibit D3b. Student Descriptive Characteristics from Analytic Sample (TFA Versus Non-TFA), Mathematics (N=42,514)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Prior Year Achievement	-0.23	-0.21	-0.03	0.92
Grade 4	0.13	0.13	-0.02	0.32
Grade 5	0.13	0.15	-0.06	0.32
Grade 6	0.22	0.27	-0.11	0.42
Grade 7	0.29	0.26	0.05	0.44
Grade 8	0.24	0.19	0.11	0.40
White	0.09	0.10	-0.07	0.27
Black	0.32	0.31	0.00	0.41
Hispanic	0.51	0.49	0.04	0.45
Asian	0.06	0.06	-0.02	0.21
American Indian	0.00	0.00	0.01	0.05
Native Hawaiian/Pacific Islander	0.00	0.00	0.00	0.05
Two or more races	0.02	0.02	-0.02	0.14
Race Not specified	0.00	0.00	0.00	0.02
Female	0.48	0.50	-0.03	0.50
EL	0.21	0.20	0.02	0.38
Special education student	0.14	0.15	-0.02	0.35
Free or reduced-price lunch eligibility	0.81	0.79	0.05	0.38
Student repeating current grade	0.01	0.00	0.01	0.07
Transient	0.10	0.09	0.01	0.25
Proportion of school year attended	0.94	0.94	-0.02	0.06

SMD = standardized mean difference; SD = standard deviation

Exhibit D4a. Teacher Descriptive Characteristics from Analytic Sample (TFA Versus Non-TFA), Reading (N=851)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Average Prior Year Student Achievement	-0.33	-0.32	-0.02	0.55
Bachelor's degree	0.89	0.85	0.12	0.25
Some postbaccalaureate	0.00	0.00	0.00	0.00
Master's degree	0.10	0.14	-0.12	0.25
Doctorate	0.00	0.00	0.00	0.02
Students of color	0.92	0.90	0.10	0.13
Female students	0.48	0.46	0.10	0.12
Special education students	0.23	0.23	-0.08	0.24
EL students	0.23	0.26	-0.13	0.21
Students eligible for free or reduced-price lunch	0.84	0.82	0.13	0.17
Number of students	50	47	0	34

SMD = standardized mean difference; SD = standard deviation

Exhibit D4b. Teacher Descriptive Characteristics from Analytic Sample (TFA Versus Non-TFA), Mathematics (N=788)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Average Prior Year Student Achievement	-0.33	-0.36	0.04	0.53
Bachelor's degree	0.91	0.84	0.19	0.26
Some postbaccalaureate	0.00	0.00	0.00	0.00
Master's degree	0.09	0.16	-0.19	0.26
Doctorate	0.00	0.00	0.00	0.00
Students of color	0.92	0.90	0.18	0.12
Female students	0.48	0.48	0.01	0.10
Special education students	0.22	0.24	-0.11	0.24
EL students	0.24	0.23	0.04	0.17
Students eligible for free or reduced-price lunch	0.83	0.81	0.17	0.17
Number of students	56	52	0	39

SMD = standardized mean difference; SD = standard deviation

Exhibit D5a. School Descriptive Demographics from Analytic Sample (TFA Versus Non-TFA), Reading (N=563)

Covariate	TFA (mean)	Non-TFA (mean)	SMD	SD
Students of color	0.90	0.89	0.07	0.14
Female students	0.48	0.48	0.04	0.05
Special education students	0.16	0.17	-0.03	0.05
EL students	0.21	0.22	-0.02	0.12
Students eligible for free or reduced-price lunch	0.79	0.77	0.14	0.16
Number of students	640	650	0	289

SMD = standardized mean difference; SD = standard deviation

Exhibit D5b. School Descriptive Demographics from Analytic Sample (TFA Versus Non-TFA), Mathematics (N=499)

Covariate	TFA	Non-TFA	SMD	SD
Students of color	0.89	0.89	0.06	0.13
Female students	0.48	0.48	-0.04	0.03
Special education students	0.16	0.16	0.04	0.04
EL students	0.21	0.22	-0.04	0.12
Students eligible for free or reduced-price lunch	0.76	0.75	0.15	0.18
Number of students	643	664	0	287

SMD = standardized mean difference; SD = standard deviation

About the American Institutes for Research

Established in 1946, with headquarters in Arlington, Virginia, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance to solve some of the most urgent challenges in the U.S. and around the world. We advance evidence in the areas of education, health, the workforce, human services, and international development to create a better, more equitable world. The AIR family of organizations now includes IMPAQ, Maher & Maher, and Kimetrica. For more information, visit AIR.ORG.

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**