

NATIONAL
CENTER *for* ANALYSIS of LONGITUDINAL DATA *in* EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington



*Examining Spillover Effects
From Teach For America
Corps Members in Miami-
Dade County Public Schools*

Michael Hansen

Ben Backes

Victoria Brady

Zeyu Xu

Examining Spillover Effects From Teach For America Corps Members in Miami-Dade County Public Schools

Michael Hansen
American Institutes for Research

Ben Backes
American Institutes for Research

Victoria Brady
American Institutes of Research

Zeyu Xu
American Institutes for Research

Contents

Acknowledgements	2
Abstract.....	3
I. Motivation and Background	4
II. TFA Placement in Miami-Dade.....	7
III. Theory of Action for TFA Spillover	9
IV. Data.....	12
V. Empirical Strategy	16
VI. Results.....	19
VII. Conclusion	25
References.....	28
Figures.....	30
Tables	32
Appendix: Data Cleaning Rules for Analysis.....	39

Acknowledgements

This research was made possible in part by generous financial support from the the John S. and James L. Knight Foundation, the National Center for the Analysis of Longitudinal Data in Education Research (CALDER), and an anonymous foundation. Research supported by CALDER is funded through Grant R305C120008 to the American Institutes for Research from the Institute of Education Sciences, U.S. Department of Education.

We gratefully acknowledge the cooperation of Miami-Dade County Public Schools, especially Gisela Feild, in providing data access, research support, and allowing us to interview district and school personnel regarding Teach For America (TFA). We acknowledge the cooperation of TFA's national and Miami regional offices in conducting this research. We thank Jeff Nelson for excellent research assistance, and helpful comments from Raegen Miller, Jason Atwood, Rachel Perera, and participants in presentations of findings at the Knight Foundation and the 2014 Association for Education Finance and Policy Conference.

The Knight Foundation funded both the TFA clustering strategy and this external evaluation of that strategy. The Knight Foundation, TFA, and Miami-Dade County Public Schools were cooperative in facilitating this evaluation and received early drafts of these findings, but none had editorial privilege over the actual contents or findings presented here. The CALDER Center has a current contract with TFA to conduct a separate evaluation of corps member impacts in a different region of the country, which TFA procured through a competitive bid process; these evaluations are independent of each other, and TFA has no editorial influence in the presentation of findings in either the current study or this separate contract.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication.

CALDER • American Institutes for Research
1000 Thomas Jefferson Street N.W., Washington, D.C. 20007
202-403-5796 • www.caldercenter.org

Examining Spillover Effects From Teach For America Corps Members in Miami-Dade County Public Schools

Michael Hansen, Ben Backes, Victoria Brady, & Zeyu Xu

CALDER Working Paper No. 113

Revised August 2015

Abstract

A sizeable body of evidence has documented the effectiveness of Teach For America (TFA) corps members at raising the mathematics test scores of their students, though little is known about the program's impact at the school level. TFA's recent placement strategy in the Miami-Dade County Public Schools, in which large numbers of TFA corps members are placed as clusters into a targeted set of disadvantaged schools, provides an opportunity to evaluate the impact of the TFA program on broader school performance. This study examines whether the influx of TFA corps members led to a spillover effect on other teachers' performance. We find that many of the schools chosen to participate in the cluster strategy experienced large subsequent gains in mathematics achievement. These gains were driven in part by the composition effect of having larger numbers of effective TFA corps members. However, we do not find any evidence that the clustering strategy led to any spillover effect on schoolwide performance. In other words, our estimates suggest that the extra student gains for TFA corps members under the clustering strategy would be equivalent to gains resulting from an alternate placement strategy in which corps members were evenly distributed across schools.

I. Motivation and Background

Teach For America (TFA) is an alternative certification program that places intensively selected recent college graduates and midcareer professionals into classrooms serving high-need students. Yet, little is known about the program's impact beyond the classrooms of individual corps members, such as whether corps members affect other teachers' classroom performance. TFA's recent placement strategy in the Miami-Dade County Public Schools (M-DCPS), in which many TFA corps members are placed as clusters into a targeted set of disadvantaged schools, provides a unique research opportunity to use the rapid expansion of TFA in targeted schools to identify the influence of corps members on colleagues' and each other's performance.

TFA operates by selecting and training corps members to teach for 2 years in high-need public schools, filling vacancies otherwise considered difficult to staff. Several prior evaluations of TFA corps members' classroom performance conclude they outperform comparison teachers in mathematics and science but perform at similar levels in reading. These evaluations come from both experimental (Clark et al., 2013; Clark et al., 2015; Glazerman, Mayer, & Decker, 2006) and quasi-experimental (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Kane, Rockoff, & Staiger, 2008; Xu, Hannaway, & Taylor, 2011) research designs.¹ The increased productivity of TFA corps members is presumed to be driven

¹ Two of these studies do not find statistically significant gains in mathematics attributable to TFA corps members: Boyd et al. (2006) and Clark et al. (2015) both estimate positive coefficients for TFA corps members in mathematics, but the full-sample estimates are not statistically significant. Both of these studies, however, produce subsample estimates that do find statistically significant gains attributed to TFA corps members. The Boyd et al. (2006) study finds that first-year TFA middle school mathematics teachers outperform other beginning teachers by about 0.05 *SD*. The Clark et al. (2015) study finds lower elementary school students (Grades PK–2) scoring 0.12 *SD* higher in reading when taught by a TFA corps member.

largely by TFA's ability to select high-quality candidates for placement in the classroom, though TFA's specific role in that selection process is still being examined in the research literature.²

The impact of TFA's placements on the broader school has not previously been examined, though TFA believes the presence of its corps members can be a catalyst for improvement, perhaps by transforming the school culture. One can reasonably hypothesize that corps members may boost the productivity of the schools in which they are placed through the *spillover effects* on other teachers—both other TFA corps members and non-TFA teachers in the school—which could affect a school's overall performance. Spillover effects may become more pronounced, in theory, as the number of TFA corps members clustered in a school increases. Jackson and Bruegmann (2009) present evidence of teacher spillovers, finding that students perform better when their teachers' peers improve over time. They find this spillover effect is especially pronounced for inexperienced teachers, which is relevant in our setting because TFA teachers are generally new to teaching.

TFA's clustering placement strategy in M-DCPS represents a unique opportunity to leverage substantial variation in TFA density within targeted schools over time to measure the association between TFA density and student achievement. In typical settings, addressing the spillover question would be difficult. Focusing on school-level spillover effects requires either across-school or within-school variation over time to identify effects. Because TFA generally targets the highest need schools in attempting to place their corps members, students' unobservable tendency to score poorly in these schools may bias spillover effects identified through across-school variation. On the other hand, using within-school variation over time requires large variation in TFA density over time that may not be present in typical situations. We therefore use within-school variation over time, generated by the

² Xu et al. (2011) estimate that the TFA effect is primarily driven by selecting candidates with high observable characteristics (selective universities, high Praxis scores, etc.), though Clark et al. (2013) find that the TFA effect cannot be explained by these differences in observables. Dobbie (2011) uses data on TFA rubrics in evaluating corps member applications and finds these measures to be predictive of performance in the classroom, independent of other observable characteristics.

clustering strategy, to measure spillover effects for this study. This has an added benefit of being arguably more exogenous than typical within-school TFA variation in the absence of clustering, because there is an outside shock in the number of available TFA corps members to hire.

This evaluation also relates to research examining school turnaround. Two years after TFA instituted the cluster placement strategy on its own, it formally partnered with M-DCPS's Education Transformation Office (ETO), the office established to oversee school turnaround efforts in the district. The ETO works with low-performing schools to help them implement one of the four federally prescribed turnaround models to improve school performance. One of these models, the *turnaround model*, prescribes low-performing schools to turnover at least 50% of the low-performing school's teaching staff. This partnership between TFA and the ETO has enabled TFA to channel corps members specifically into schools that were labeled as chronic low performers. Most of the schools that received a large influx of corps members—including all middle schools and high schools—implemented the *turnaround model*. Although this partnership may aid schools in building committed staffs in short order, prior evidence of success for such a strategy is generally weak overall, as described in the Institute of Education Science's *Practice Guide* on the topic (Herman et al., 2008). However, a more recent, rigorous study of school turnaround efforts in California points to the turnaround model as having the largest associated effect of the four federally prescribed models (Dee, 2012), further supporting the rationale behind these models. How well this strategy of using clusters of TFA corps members to help fill staffing needs in these turnaround schools is an open question that is germane to our inquiry here.

This study investigates the role of TFA spillover on both non-TFA and other TFA colleagues in contributing to changes in student learning in targeted schools. We use administrative data from M-DCPS for the 6 school years between 2008–09 and 2013–14, with school fixed effects to leverage the variation in TFA density over time within schools. In addition to the administrative data, the study team also conducted live semi-structured interviews with district personnel who have worked with the TFA

program and school principals or assistant principals in schools where TFA corps members have been placed since the 2009–10 school year. Though an analysis of this qualitative data is beyond the scope of this paper (see Hansen, Backes, & Nelson, 2015), we supplement the empirical analysis presented here with interview responses, where appropriate. In the summary of our findings, we observe that many of the schools chosen to participate in the cluster strategy experienced large subsequent gains in mathematics achievement. Our results indicate these gains were driven in part by the composition effect of having larger numbers of effective TFA corps members. However, we do not find any evidence that the clustering strategy led to any substantive spillover effects on schoolwide performance.

II. TFA Placement in Miami-Dade

TFA started placing corps members in M-DCPS in 2003, with 35 initial placements.³ During the early period of TFA’s presence in the district, the placement of corps members in schools did not adhere to an overarching strategy, except for TFA’s requirement of placing corps members in schools where 70% or more of students are eligible for free or reduced-price lunch (FRL). For all schools meeting this criterion, corps members were placed wherever TFA could establish sufficient rapport with school principals as to allow them to be considered for vacancies. This approach to placement resulted in TFA corps members being spread thinly across many schools in the district. By the summer of 2008, TFA’s yearly cohort size was approaching 50 corps members, resulting in a total presence of 90 active corps members (representing two cohorts) staffed across 48 schools during the following school year.

Beginning with the 2009–10 school year, TFA began a clustering strategy in which new placements were purposely assigned to target schools within designated high-need communities. TFA’s clustering placement strategy grew out of an interest in accelerating TFA’s impact on student outcomes

³ This section draws heavily on conversations with the TFA Miami regional office as well as those with personnel in the M-DCPS central office. We thank them for generously providing details of the program.

specifically in these communities. Based on conversations with those originally involved in the design of the clustering strategy, these accelerated outcomes were hypothesized to be achieved through several means, three of which are relevant for spillover.⁴ First, concentrating TFA corps members should improve student outcomes as a mechanical result of staffing a greater quantity of high-performing teachers in target schools (the composition effect). Second, TFA believed a critical mass of young, energetic corps members would possibly spillover into non-TFA teachers' classrooms, and potentially affect the whole school (the spillover effect on non-TFA). And third, TFA expected placing multiple corps members in the same schools would increase corps members' sense of support and satisfaction from the program, which was hoped to lead to better performance among active corps members (the spillover effect on other TFA). Beyond its benefits to students, TFA expected that the strategy of higher concentrations of corps members in fewer schools would be beneficial from a management perspective: TFA could better manage and build deeper relationships with building-level administrators and provide in-person support to corps members more efficiently. The new placement strategy was conceived by the regional TFA office located in Miami, endorsed by M-DCPS, and encouraged with external funding. Since being implemented in M-DCPS, this placement strategy has been loosely replicated in other, mostly rural, TFA regions.

The TFA clustering strategy soon became one piece in a larger school turnaround effort of M-DCPS's ETO, which was established in 2010 to administer turnaround efforts in schools designated as the persistently lowest achieving schools in the district. Due to the natural overlap in the targeted schools for both TFA's clustering strategy and the ETO, they partnered to start placing TFA teachers in select ETO schools starting in the 2011–12 school year. Both entities viewed this as a mutually beneficial

⁴ In addition to the three hypotheses named in the text, TFA also hypothesized corps members would be retained in placement schools beyond their commitment period at higher rates, and students in targeted schools would be exposed to multiple TFA corps members during normal matriculation, thus potentially making a cumulative effect on the high-need student populations. Follow-up studies in this project evaluate the clustering strategy's association with these outcomes (see Hansen & Backes, 2015b; Hansen, Backes, & Brady, 2015).

partnership—the ETO valued the flow of corps members to vacancies that are otherwise difficult to staff,⁵ while TFA viewed this as a way to strategically target their efforts in the highest need schools, which was expected to maximize their impact on students. This partnership further accelerated the growth of the total number of corps members working in the district.

The growth of the TFA corps and its density is readily apparent in the placement numbers during the 6 school years of the data used for this analysis. Table 1 presents TFA corps member assignment figures over time. In the 2008–09 school year, the year immediately preceding the clustering strategy, there was an average of slightly less than two TFA corps members in each school where they were hired. In the years following, the number of schools containing any TFA corps members dropped by about half and the number of active TFA corps members in the district more than tripled, resulting in an average of nearly 10 corps members per school where there was any presence.⁶ The net result was a jump in the proportion of TFA corps members in placement schools, going from 2% to 4% in 2008–09 to 14% to 18% in 2012–13.

III. Theory of Action for TFA Spillover

The theory behind spillover presumes that all teachers affect each other: More effective peers, on average, will promote effectiveness in colleagues.⁷ Jackson and Bruegmann (2009) present evidence

⁵ TFA placements under the clustering strategy are still heavily dependent upon position vacancies and principal buy-in. Neither schools nor the district made explicit decisions to fill a certain number of vacancies with TFA teachers, but considered the pool of eligible incoming TFA candidates to fill vacancies in target schools. The ultimate decision to hire a TFA corps member was left to school principals, though under the clustering strategy, principals had to hire at least two TFA corps members in the same school. Thus, the density of TFA corps members in a school was determined by available vacancies, principals' selection of corps members to fill them, and the size of the incoming TFA corps cohort eligible for placement. There was no strategy of targeting a certain number of TFA corps members in each school (aside from the two corps member minimum).

⁶ When conducting interviews with school and district administrators, one concern we heard about the large increase in TFA placements was a potential dilution of the quality of the TFA pool due to the district filling about three times as many placements with TFA corps members. Any decrease in quality could also be a confounding factor in our attempt to identify spillover effects. We address this issue further in Section VI below.

⁷ A second type of teacher spillover has been addressed in prior studies, which model the influence of teachers in one subject on students' other tested subjects (though taught by other teachers). These are primarily addressed as

of productivity spillovers among elementary school teachers in North Carolina, although they do not investigate the mechanisms of spillover. Presumably, spillover may arise through sharing instructional resources, coaching each other, or simply motivating each other. The authors find novice teachers are particularly responsive to the arrival of effective colleagues, but it is not clear that novices are particularly influential in their ability to shape others' performances. Applying these findings to TFA, where almost all of these corps members are novices, one may suspect that the density of TFA corps members might have the strongest effect on other TFA corps members, and only a moderate effect on more experienced non-TFA teachers. Consequently, as we look for evidence of spillover effects from TFA corps members, we attempt to differentially identify spillover that may affect these two groups.

During in-person interviews, multiple school principals expressed opportunity for TFA corps members to influence their colleagues' practice, suggesting spillover could be plausible. The most common opportunity principals cited is Common Planning, a districtwide effort to organize teachers of the same grade (in elementary grades) or subject (in secondary grades) in a regular shared planning period (usually weekly) to coordinate efforts and promote mutual learning. One elementary school principal related, "Through Common Planning, the sharing of best practices, through our data checks, TFA corps members will show what worked—their determination spread, it's contagious." In addition, clustered TFA corps members may also influence other teachers across the school more broadly, which was supported by a district administrator, "When you have a few corps members in the same school, they tend to feed off of each other, and often times there's a big change in the school culture." Several principals participating in the interviews indicated specific mechanisms through which these cultural changes may occur, including TFA corps members' high energy levels, high expectations, and outreach to parents, all of which they cited as being noticed and modeled by other teachers. Thus, if these

causes for concern in estimating teachers' value-added contributions to student learning (e.g., Koedel, 2009; Yuan, 2015) and are beyond the scope of our inquiry here.

behavioral spillovers translate into student achievement, a TFA spillover effect among all teachers broadly may be plausible, in spite of TFA corps members' relative inexperience.⁸

Finally, we note that this study looks for evidence of fairly large spillovers associated with the concentration of TFA corps members. The estimated magnitude of productivity spillovers across teachers is relatively small in the Jackson and Bruegmann (2009) study. They estimate that an increase of one standard deviation (*SD*) in the mean estimated value-added of a teacher's peers is associated with an increase of 0.0398 *SD* in mathematics test scores. In our data, the average TFA teacher is about 40% of a *SD* more effective in mathematics than the average replacement teacher in teacher value added (authors' calculation, ignoring experience differentials), so the equation for the expected increase in student test scores associated with replacing an average teacher with a TFA teacher would be the following:

$$\frac{1}{n}(0.0398)(0.40),$$

where *n* represents the number of teachers in the peer group (Jackson and Bruegmann consider other teachers within a grade and school as peers in their sample of elementary school students). Thus, for a grade with five teachers, the expected increase in other students' test scores due to replacing one with a TFA corps member would be 0.003 *SD*, an effect that is too small to detect, given the standard errors in most specifications using the available data. For our study, however, we are not concerned with confirming a TFA spillover effect comparable in magnitude to these earlier estimates; rather, we wish to explore whether TFA produces spillovers on both TFA and non-TFA teacher colleagues are large enough to make a significant difference in a school's performance, even when the change in staffing

⁸ Using Australian data, Bradley, Green, and Leeves (2007) find evidence of behavioral spillovers in teacher absences associated with the arrival of teacher colleagues that show prior patterns of high or low absences. The authors do not provide any evidence of this behavior's effect on student achievement.

may be modest. The absence of such large effects would weaken the justification for clustering together corps members under this strategy.

IV. Data

We use detailed student-level administrative data that cover M-DCPS students linked to their teachers for 6 school years (2008–09 through 2013–14). M-DCPS is the largest school district in Florida and the fourth largest in the United States. The district has large populations of minority and disadvantaged students, typical of regions TFA has historically targeted; about 60% of its students are Hispanic, 30% Black, and 10% White, and more than 60% of students qualify for FRL.

The student-level longitudinal data we use in the analysis contain reading and mathematics scores on the Florida Comprehensive Achievement Test (FCAT).⁹ Students' FCAT scale scores are converted to z-scores based on the mean and *SD* for that particular subject-grade test in the M-DCPS sample. Test scores in each year are outcomes, and prior-year test scores are used as controlling covariates in the value-added approach when estimating student outcomes used in the analysis (described further below); only students with valid pretest scores are included in the analysis sample. In addition to standardized test scores, we observe a variety of student characteristics: race, gender, FRL eligibility, limited English proficiency (LEP) status, whether a student is flagged as having a mental, physical, or emotional disability, attendance, and disciplinary incidents. In addition, all students are linked to teachers through data files that contain information on course membership.¹⁰

⁹ From the 2008–09 school year through the 2010–11 school year, all students Grades 3–10 took the FCAT in both mathematics and reading. However, with the introduction of End-of-Course (EOC) exams in 2011–12, the mathematics portion of the FCAT will only be administered to Grades 3–8 from 2011–12 forward. For students taking an EOC exam in 2011–12 through 2013–14 (for example, Algebra I), we consider their previous year's FCAT score to be their lagged test score. See appendix for more information.

¹⁰ Teachers of record in students' core mathematics and reading courses are linked to them for the analysis. Student observations linked to multiple teachers (e.g., due to coteaching, student mobility) are weighted in proportion to the amount of time spent with each teacher, based on available enrollment data. Please see appendix for more information.

Teacher personnel files in the M-DCPS data contain information on teachers' experience levels, education attainment, demographics, and other supplemental background variables. These will likewise be used as covariates for various models in the analysis that follows. One variable included in the data is whether the teacher is a TFA corps member (both active corps members during their two-year commitment or former corps members now considered TFA alumni). Given the importance of this variable in the analysis, we externally validated this variable with corps member lists from TFA, and found nearly perfect overlap between the district-supplied variable and TFA lists (any person found on either list is flagged as TFA). Note that TFA in the data refers to all TFA-affiliated teachers, including both active corps members and alumni who continue to teach in M-DCPS beyond their 2-year commitment.

Table 2 presents descriptive statistics of the two samples utilized for the study (one for each subject: mathematics and reading). The table groups schools within each sample separately by TFA cluster status, which we define as any school in which two or more corps members from the same cohort are placed, starting in the summer of 2009 and after. The samples used are, of necessity, limited to grades and subjects in which standardized tests are administered to students. Hence, the few schools in which all TFA corps members are placed outside of these grades and subjects are not flagged in the cluster TFA subsample.

As shown in Table 2, the cluster schools where TFA corps members have been placed since 2009 tend to be very observationally dissimilar to the rest of the district. About two thirds of students in noncluster schools are Hispanic, and more than three fourths of students in cluster schools are Black (based on student characteristics in the analysis sample in tested grades and subjects). In addition, the share of FRL-eligible students is about 20 percentage points higher in placement schools. This is consistent with TFA placement patterns of choosing high-need schools in which to place its corps members. In addition, student achievement on the FCAT in cluster schools is about 0.6 *SD* and 0.5 *SD* lower in reading and mathematics, respectively.

Differences also emerge with observable teacher characteristics, although they are not as stark as the differences among students. Teachers in noncluster schools are about five percentage points more likely to have at least a master's degree, and average an additional 2 to 3 years of experience. By construction, the share of TFA teachers is much higher in the cluster sample.¹¹ Also, teachers in placement schools are significantly more likely to be Black and less likely to be Hispanic, relative to the noncluster schools.

Because TFA corps members and alumni are the primary focus of this study, it is helpful to examine descriptive statistics of these placements over time for those appearing in the analysis sample. Table 3 reports corps member placements over time included in the analysis sample (reporting information parallel to Table 1) and also presents descriptive statistics of the classrooms TFA teachers are leading. Note that the TFA proportion values and descriptive statistics include only active corps members; however, given the high attrition of TFA corps members out of the classroom after 2 years, these figures only slightly vary when reporting on active corps members and TFA alumni combined. For instance, cluster schools averaged less than one TFA alumni remaining in the school beyond the initial 2-year commitment.

Two particular elements of Table 3 are worth highlighting. First, the schools with any TFA corps members that are included in the analysis sample tend to have even higher percentages of TFA in them (comparing against the percentages reported in the entire district in Table 1 above). These figures indicate TFA corps members are overrepresented among tested grades and subjects, which is unsurprising because TFA teachers are most commonly granted a provisional license to teach in core academic subjects (mathematics, science, and English language arts) rather than untested subjects (e.g., history, art) that are omitted from the analysis sample. When conducting interviews, we did not find any

¹¹ There are some TFA corps members and TFA alumni in the noncluster sample due to residual TFA corps members left over from before the cluster period.

evidence of principals systematically sorting TFA corps members to a teaching assignment that would be tested (conditional on the corps members' license areas).¹²

Second, the table shows a large and notable jump in mathematics test scores (the difference between the pretest and current mathematics achievement scores) among students taught by TFA corps members. The posttest scores jump considerably in the last 3 years of data, an increase of well 0.20 *SD* of student achievement, whereas the increase in pretest scores during that period is much smaller in magnitude. This jump in performance during these last 2 years is particularly noteworthy for two reasons: First, this jump in performance coincides with the largest single year-to-year increase in the total number of TFA corps members in the district (84 corps members; see Table 1); and second, it also coincides with the initiation of TFA's formal partnership with the district's ETO to help turnaround low-performing schools. These two coincident events could potentially cloud our ability to identify a TFA spillover effect, as concurrent schoolwide turnaround interventions will be confounded with the spillover effect if they are correlated with high-dosage TFA schools.

This possible bias prompts us to inspect the performance trajectories of TFA cluster schools with those of noncluster ETO schools; these are presented in Figure 1. This figure shows that this large increase in mathematics test scores appears to be common among both groups of schools; however, the surge in mathematics test scores for TFA cluster schools appears to begin 1 year earlier in the cluster schools and is larger in magnitude than that observed among the remaining ETO schools. No apparent improvement is observed in reading test scores in either group; both groups have shown declines relative to their 2008–09 performance. Note that most, but not all, of the 37 TFA cluster schools are also considered ETO schools, whereas the noncluster ETO group contains 28 schools.¹³ Hence, to avoid

¹² Principal responses were mixed on this question, with some describing an inclination to staff TFA corps members specifically in tested grade and subject assignments, others describing a strategy that kept TFA out of these assignments for at least a year, and a third group reporting no particular strategy.

¹³ The list of schools considered ETO by the district has grown over the last several years; we identify a school as

attributing this rise in achievement solely to the TFA clustering strategy, we will control for year-specific ETO effects in our analysis.¹⁴

V. Empirical Strategy

To model the influence of TFA's clustering strategy on overall school performance, we begin with a straightforward value-added regression predicting student achievement for student i in school s in classroom c at time t on test scores (A_{ist}) as a function of prior student achievement (A_{it-1}), student characteristics (X_{it}), classroom characteristics (X_{ct}), and a school fixed effect (γ_s).¹⁵ Studies of TFA effectiveness generally estimate an equation similar to the following:¹⁶

$$A_{ist} = \alpha A_{it-1} + \beta_1 X_{it} + \beta_2 X_{ct} + \gamma_s + \beta_3 TFA_{it} + \varepsilon_{ist}. \quad (1)$$

The spillover effects of interest for this paper deal with TFA teachers affecting others, but how is that best empirically modeled? To address this question, we need to determine how to quantify the density of TFA teachers in schools. There are two dimensions to this measurement that must be considered: (1) Through what group are spillover effects transmitted? And (2), how is the composition of TFA teachers measured in the group? Both of these dimensions are discussed in turn below.

ETO if it has ever been considered an ETO. The TFA cluster group in Figure 1 includes all 37 TFA cluster schools, regardless of ETO status.

¹⁴ When not controlling for ETO time trends, some estimates find a positive and significant effect on math test scores associated with increasing the TFA density within a school. As shown below, the regressions including ETO controls generally do not find TFA density to have a statistically significant effect on achievement.

¹⁵ The vector of prior-year test scores contains cubic functions of prior test scores in both reading and mathematics. The vector of student characteristics includes the following: race, gender, FRL eligibility, LEP status, and mental, physical, or emotional disability status. The vector of classroom characteristics includes class size, classroom-level averages of prior-year test scores, and classroom-level averages of each of the student characteristics listed above. Teacher controls include teacher race, gender, experience, and whether the race of the teacher matches that of the student. The student characteristic, class average, and teacher demographic controls are interacted with grade indicator variables to allow differences in the influence of these variables across grades. The estimating equation additionally includes indicator variables on grades and years.

¹⁶ Studies using this approach include Boyd et al. (2006); Clark et al. (2013); Glazer et al. (2006); and Kane et al. (2008).

First, how are spillover effects transmitted? The clustering strategy, as implemented in M-DCPS, focused on increasing the presence of TFA corps members in schools generally, so considering all teachers within a school as the relevant peer group would be a natural way to approach this problem. Yet, defining the peer group this way implies that TFA corps members' influence is broad, potentially reaching others who do not share similar grade or subject assignments. If clusters of TFA affect the school culture in such a way as to promote greater productivity overall (e.g., due to higher student expectations or motivating other teachers to exert more effort), then a parameterization that defines a teacher's peers broadly should pick up this type of spillover. Alternatively, spillover effects may be more concentrated. In select schools, TFA has worked with principals to stack a particular department (typically mathematics or science) with TFA corps members. Learning among teachers during Common Planning may constitute the most tangible places for TFA corps members to affect others' practices. In these cases, we would expect that a parameterization that defines the peer group more narrowly would more likely identify the spillover effect. In the analysis that follows, we estimate models using two separate definitions of the relevant peer group—the first defines peers as any colleague within the school, the second defines peers as any colleague within the same grade (for elementary school teachers) or same subject (for middle and high school teachers). We prefer, though, the grade- or subject-level peer specification, because this seems the most plausible avenue for the transmission of spillover effects and is also the level at which Jackson and Bruegmann (2009) identify their effects. As shown below, results are mostly similar across these two definitions.

Second, how is the composition of TFA teachers measured in the peer group? Again, there is no clear answer as to how to measure the density of TFA peer members for a given teacher—one could either directly count all TFA teachers with the peer group or convert that number to represent the percentage of peer teachers who are affiliated with TFA. Alternatively, if a critical mass of TFA is needed as a catalyst for transformation, a threshold-based approach to identifying spillover may be more

appropriate. Given the ambiguity on how to quantify this variable, we present the results from both count and percentage specifications, but consider the percentage metric our preferred specification because it most closely mirrors the weighted average effectiveness measure that Jackson and Bruegmann (2009) use.

To provide a description of the variation in this TFA density in the data, Figure 2 presents a school year–level histogram of TFA density (measured as a percentage of TFA among instructional staff in the school), for all observations in which at least one TFA is present. As pictured, the modal school-year observation has fewer than 5% of teachers who are TFA (among those with nonzero TFA percentages). Less than half of all unique school-year observations have TFA density exceeding 10%. Roughly 40 observations show particularly high levels of TFA density, in which more than 20% of a school’s staff are affiliated with TFA.

Next, we include these various measures of the TFA corps members’ concentration in a school ($TFA_DENSITY_{st}$) that interacted with the indicator for TFA corps member, as shown below:

$$A_{ist} = \alpha A_{it-1} + \beta_1 X_{it} + \beta_2 X_{ct} + \gamma_s + \beta_3 TFA_{it} + \beta_4 TFA_DENSITY_{st} + \beta_5 TFA_DENSITY_{st} * TFA_{it} + \varepsilon_{ist}. \quad (2)$$

In the equation above, $TFA_DENSITY$ is a measure of TFA density. These measures of the TFA corps members are either count or percentage variables in the relevant peer group. These parameters are intended to capture any differential in outcomes that may be associated with differences in the composition of the corps members within the implementation of the cluster strategy. The coefficient estimate on TFA density ($\widehat{\beta}_4$) provides evidence of spillover among all non-TFA teachers generally, and the interacted TFA density * TFA coefficient ($\widehat{\beta}_5$) captures the extent to which the spillover effects differentially accrue to TFA teachers.¹⁷ Note the inclusion of school fixed effects implies that the variation in schools’ densities of TFA teachers over time is driving the resulting estimates.

¹⁷ To allow for students with multiple teachers, regressions are run using the Full Roster Method (Hock & Isenberg,

Equation 2 is a variation of a difference-in-difference (DD) design. A more conventional DD design would use a binary variable on cluster and noncluster schools rather than the continuous TFA density variable employed here. We choose to control for TFA density directly to distinguish between relatively high- and low-density contexts because even within cluster schools, TFA density varies considerably both across cluster schools and over time.¹⁸

VI. Results

Estimating the Baseline Effects of TFA Corps Members on Their Students

To compare the TFA corps members in our study with previously published research, we display the results of the basic teacher value-added regression represented by Equation 1 in Table 4. In all of the results that follow, all grades are combined into a single regression, with grade-specific intercepts and slopes for relevant control variables. Controls include cubic terms in a student's previous mathematics and reading scores, race, gender, and FRL status (all interacted with grade level). In addition, we take class averages of these demographic variables and interact them with grade level. These estimates, therefore, can be interpreted to represent the contributions of TFA teachers over and above the average teacher with similar student backgrounds and test scores. Also interacted with grade level are teacher characteristics (experience, race, race-matching with students, gender, and degree attainment) and class size.¹⁹ Finally, we cluster standard errors at the school level.

The first column shows a basic OLS regression while controlling for TFA and teacher experience separately. For mathematics, consistent with most prior studies, the TFA effect is positive and statistically significant. We also find a null effect for reading scores in our OLS regressions, shown in

2012), where observations are at the student-teacher link level, and are weighted differentially by teacher dosage. Please see appendix for more details.

¹⁸ Prior studies have used similar specifications to directly control for treatment intensity (e.g., Draca et al., 2011).

¹⁹ Teachers with missing values for experience are coded as 0 and are flagged with a missing experience indicator variable, which is included as a control in the regressions.

column 5. We next investigate whether TFA corps members have differential returns to experience by interacting TFA and years of experience (columns 2 and 6), and generally do not find a differential return to experience. In other words, both TFA and non-TFA teachers appear to become more effective with experience at similar rates. Columns 3 and 4 (mathematics) and columns 7 and 8 (reading) add school fixed effects, and the point estimates increase for all specifications. This is presumably due to the TFA coefficient being downwardly biased in OLS because corps members are placed in relatively disadvantaged schools; failing to add the school fixed effects attributes the school's low performance trajectory to the corps members rather than to school-specific unobservable performance. The increase in the point estimates in reading is large enough that TFA corps members now significantly outperform non-TFA colleagues by a modest 0.02 *SD*.

The magnitude of our TFA mathematics effect—about 10% of a *SD* of student learning on standardized test scores—falls roughly in the middle of previous studies. Relative to the papers discussed earlier, our estimate is somewhat smaller than Glazerman et al. (0.15 *SD*) and Xu et al. (0.13 *SD*), and larger than Clark et al. (0.07 *SD*), Kane et al. (0.02 *SD*), and Boyd et al. (no effect). Because the impact of a 0.10 *SD* improvement in test scores varies across school grades (representing approximately 20% of a school year in Grade 4 and 40% of a school year in Grade 10), we convert grade-specific TFA effects to months of student learning using the average annual gain estimates reported in Hill et al. (2008).²⁰ After converting, the weighted average TFA effects equate to a 34% boost in learning beyond average annual student gains in mathematics. This effect is equivalent to 3.4 months of learning, based off of a 10-month school year, relative to the average student assigned to a non-TFA teacher in the same school.

²⁰ When estimating grade-specific coefficients for the TFA variable, the standard errors increase substantially, and most are not statistically distinguishable from other grades' coefficients or zero. For brevity, we do not report them here.

The significance of the estimated TFA reading effect is a notable departure from the prior literature on TFA teachers' effectiveness, which generally shows no significant differences in reading (Clark et al., 2015, excluded). Although the point estimate is a modest 0.022 *SD*, comparing this against the magnitudes of the estimates from greater experience (in columns 7 and 8) shows the TFA effect to be roughly equal in magnitude to the effect of having much more senior teachers in the classroom in this data. In other words, though the estimates are modest in size, they imply a meaningful increase in student learning.

Estimating the Spillover Effects of TFA Onto Colleagues' Students

We next turn to the main research question of this paper: whether the density of TFA members in a school leads to a measureable change in student achievement beyond their own classrooms. Regression results incorporating the parameters capturing spillover are presented in two tables: Table 5 reports the results from specifications where the peer group is composed of all teachers in the school, and Table 6 presents results where peers are composed of those in either the same elementary grade or secondary subject. Mathematics results are presented in columns 1 and 2 of each table, and reading results are in columns 3 and 4. TFA density is measured as a count variable in columns 1 and 3 of each table, and as a percentage of peers in columns 2 and 4. As described above, the percentage specification is our preferred specification.

Focusing first on the coefficient estimates on TFA density, we see no consistent evidence of spillovers to non-TFA teachers from the clustering of TFA teachers. Seven of the eight reported TFA density coefficients across Tables 5 and 6 show no significant difference from zero. Though the count specification in mathematics is positive and statistically significant in Table 5, note that the main point estimate from TFA status is considerably lower in this particular specification (compared to the other point estimates of TFA on mathematics in Tables 4, 5, and 6); thus, this specification trades off a lower main TFA effect with a greater positive weight on TFA density. We additionally estimated a series of

specifications (not reported here for brevity) that modeled TFA density as a threshold at different values, and virtually all of them showed qualitatively similar null effects of TFA density; hence, we found no evidence supporting the “critical mass” hypothesis for cultural transformation.

Recall that the point estimates on the interacted TFA * TFA density variables represent the performance differential for TFA teachers in schools with increasing concentrations of TFA. Again, most of the coefficients reported across Tables 5 and 6 are not significantly different from zero, suggesting no significant spillover effects. Those specifications that are significantly different are from the less-preferred count specifications. Our critical mass threshold explorations turned up no significant differences on these interacted variables either. We conclude that the evidence favors no effect of TFA density on TFA corps members.

Examining the Potential Tradeoff Between Quality and Quantity in Clustering

As a final analysis, we investigate whether the increase in the quantity of TFA corps members placed in the clustering strategy was accompanied by a decrease in the classroom productivity of these corps members. This inquiry is motivated by the concern that the increase in TFA density over time (which we are using to identify spillover effects), if concurrent with a decrease in quality, may understate actual spillover, because the empirical model estimates a single TFA effect over the entire period.

Not only is this possible drop in quality an empirical concern, but it is also a practical concern: When conducting interviews with district administrators, this concern was the primary one voiced about the clustering strategy and was mentioned by all administrators interviewed. One district administrator

stated, “I would revisit the way they are recruiting TFA; I think everyone recognizes that when you [have grown this much], there is obviously a danger in loss of quality.”²¹

Table 7 presents the results of our investigation into changes in TFA effect estimates over time. These results are produced by re-estimating the main TFA effects as detailed in Equation 1, and then adding either cohort-specific (on the left) or year-specific (on the right) interaction terms with the TFA indicator variable. If quality is declining with more recent placements, we expect to see a trend of negative point estimates on the interaction variables representing these placements.

The cohort-specific point estimates show a moderate level of fluctuation over time, but no clear downward trend with later cohorts.²² The year-specific point estimates fluctuate less (as more corps members are included in each grouping), but again do not demonstrate a downward trend. If anything, performance in reading appears to be improving in later cohorts and years, though point estimates are only marginally significant. Based on these results, we conclude the spillover effects estimated in Tables 5 and 6 are unlikely to be tainted by a concurrent drop in productivity among TFA estimates.²³

Discussion of Findings and Limitations

In sum, the results presented in Tables 5 and 6 do not find a consistent pattern of TFA density affecting student achievement, and our preferred specification estimates (columns 2 and 4 of Table 6) are uniformly not significantly different from zero, suggesting no consequential TFA spillover effect in the data. Based on Table 7, we have no reason to believe a concurrent decrease in quality deflates

²¹ Further discussion of the perceived quantity-quality tradeoff among district administrators is presented in Hansen, Backes, & Nelson (2015).

²² For the 2007 cohort, the data do not include their first year of teaching, and there are very few observations, likely leading to the imprecise estimates found.

²³ Though we do not find any clear empirical evidence of lower classroom productivity among more recent placements, this does not necessarily imply all TFA placements in the district are of the same quality over time. To be included in the analysis sample, a teacher must be assigned to a tested grade and subject, and teach in this assignment for the full school year. If selection into these tested classroom assignments or the premature attrition of TFA corps members has changed during this period, the analysis in Table 7 will not detect them. Hence, although we use the results of Table 7 to remove the possibility of lower productivity confounding our spillover estimates, it should not be interpreted as definite evidence on the current health of the TFA corps in the district overall.

spillover that may be otherwise present. However, there may be some other mitigating factors behind the interpretation of these results; we briefly discuss these here.

The first is related to our ability to detect an effect. As described previously, the spillover estimates in the Jackson and Bruegmann (2009) study are very small—our hypothetical of replacing one teacher with a TFA corps member in a grade with five teachers equated to a 20% change in TFA density and an improvement of 0.003 *SD* in peer teachers' student test scores. Our estimates are precise enough to detect effect sizes as small as 0.04 *SD* for the same 20% change in TFA density. In light of this minimum detectable effect, it is no surprise that we cannot detect a spillover effect in reading, where the main effect of having a TFA teacher on their own students' test scores is 0.02 *SD* (presumably, any spillover effects onto other teachers should be only a fraction of the main TFA effects on their own students). In mathematics, however, the main TFA effect exceeds 0.10 *SD*, and an effect size of 0.04 *SD* is roughly equal to the performance advantage of a teacher with 1 year of prior experience compared to a novice, based on the point estimate in column 3 of Table 4. The minimum detectable effects in our estimates are small enough to rule out substantial changes in test score performance of colleagues' classes as a result of TFA placements. Given the hypothesized influence of TFA as a catalyst in promoting whole-school improvement, the results here suggest any TFA spillover effect is too small to systematically promote such a change, particularly when the typical "intervention" consists of filling a few vacancies with TFA corps members.

A second mitigating factor is our exclusive focus on student test scores in this study. Spillover effects from TFA corps members onto other teachers may be transmitted through a variety of behaviors, which may affect students, other teachers, or the school culture in many ways. Although we do not find any evidence of meaningfully large gains on test scores, this does not remove the possibility of changes conveyed through these other mechanisms. We have produced a series of related studies evaluating various outcomes in the context of TFA's clustering strategy in M-DCPS, including teacher mobility and

retention of TFA and non-TFA colleagues (Hansen, Backes, & Brady, 2015), the cumulative effects of student exposure to TFA in cluster schools (Hansen & Backes, 2015a), and TFA impacts on nontested student outcomes (Hansen & Backes, 2015b). Further, our analysis of the qualitative data collected during our semi-structured interviews with school and district administrators provides additional nuance to the influence and the limitations of TFA corps members in these schools (Hansen, Backes, & Nelson, 2015). Summarizing across these studies, it appears that the composition effect of TFA in targeted cluster schools is the most consequential in improving learning in those schools: The increase of TFA cohort size over time has allowed more students to be exposed to classrooms taught by TFA corps members. Other hypothesized outcomes, including improved retention of corps members, lasting cultural change, and cumulative student learning gains show little evidence of making significant contributions to observed gains.

VII. Conclusion

The research questions motivating this study ask whether TFA corps members affect other teachers in their school—both non-TFA teachers and other TFA corps members—through spillover effects. We exploit the variation in TFA corps member densities within schools over time, which occurred due to the implementation of the TFA clustering placement strategy in M-DCPS, to investigate these questions. With student-teacher linked administrative data from M-DCPS, we estimate changes in teacher effectiveness in reading and mathematics that are associated with changes in TFA teacher densities using a school fixed effects model.

In summary, we find little evidence of meaningful spillover effects from TFA corps members on student test scores in the short term—neither on non-TFA teachers nor on TFA corps members in the same schools. We explored a variety of specifications of the TFA density measure, and virtually all resulted in no significant differences resulting from these changes induced by clustering. However, we

do find robust evidence of TFA effects on mathematics test scores in the range of 10% of a *SD* of student achievement, or averaging over 3 months of learning. This is also the first study to document TFA teachers outperforming comparison teachers in reading, by an estimated 0.02 *SD* of student learning.²⁴

Was the cluster placement strategy a success in M-DCPS? It may be, in spite of the lack of spillover. TFA stated two primary objectives in designing and implementing the cluster placement strategy: (1) to accelerate TFA's influence in student outcomes in particularly disadvantaged settings and (2) to provide more support for TFA corps members through an increased presence in schools and in the district overall. Although spillover was an expected result of the strategy, it was not a primary objective. Given the observed patterns of corps member placement in recent years, it is clear that TFA's presence in the district has substantially increased, and the presence of TFA in some of the highest need schools in the district has likewise increased. Thus, the composition effect alone—where vacancies in high-need schools are filled with relatively effective TFA corps members—implies that TFA's increased presence has made a significant difference on student learning in the district. In this way, the clustering strategy has at least partially achieved its objectives. Furthermore, we do not find any evidence that the large increase in the number of TFA placements in recent years was associated with a reduction in TFA effectiveness.

The results here, however, provide no evidence of spillover on student test scores in the short term. In other words, there is no reason to expect that the extra student gains for TFA corps members under the clustering strategy would be any different (in the aggregate) than the gains that could result from an alternate placement strategy where corps members are more evenly distributed across schools. Yet, even if the placement strategy does not affect teacher spillover, how teachers are placed across schools will affect districtwide achievement gaps—broad placement of TFA corps members will boost

²⁴ Clark et al. (2015) find that TFA teachers out-perform non-TFA teachers in a subsample of prekindergarten through grade 2 teachers, but not in other subsamples.

many students' mathematics performance slightly, whereas focusing on high-need schools boosts student achievement in mathematics in a more targeted way. By focusing these placement efforts in some of the most disadvantaged and low-performing schools in the district rather than spreading corps members broadly across many schools, the clustering strategy has accelerated growth in schools that are in the greatest need, and within-district achievement gaps are likely reduced (albeit very modestly) as a result.

References

- Boyd, D., 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.
- Bradley, S., Green, C., & Leever, G. (2007). Worker absence and shirking: Evidence from matched teacher-school data. *Labour Economics, 14*(3), 319–334.
- Clark, M., Chiang, H., McConnell, S., Silva, T., Sonnenfeld, K., Erbe, A., . . . Puma, M. (2013). *The effectiveness of secondary math teachers from Teach For America and the Teaching Fellows programs* (No. 7890). Princeton, NJ: Mathematica Policy Research.
- Clark, M., Isenberg, E., Liu, A., Makowski, Z., & Zukiewicz, M. (2015). *Impacts of the Teach For America Investing in Innovation Scale-Up*. Washington, DC: Mathematica Policy Research.
- Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus* (Working Paper No. 17990). Washington, DC: National Bureau of Economic Research.
- Dobbie, W. (2011). *Teacher characteristics and student achievement: Evidence from Teach For America* (Unpublished Manuscript). Cambridge, MA: Harvard University.
- Draca, M., Machin, S., & Van Reenen, J. (2011). Minimum wages and firm profitability. *American Economic Journal: Applied Economics, 3*(1), 129–151.
- 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.
- Hansen, M., & Backes, B. (2015a). *Teach For America impact estimates on nontested student outcomes* (Working Paper #142, Forthcoming). Washington, DC: CALDER Center.
- Hansen, M., & Backes, B. (2015b). *Persistence of TFA learning effects* (Unpublished Manuscript). Washington, DC: CALDER Center.
- Hansen, M., Backes, B., & Brady, V. (2015). *Teacher attrition and mobility during the Teach For America clustering strategy in M-DCPS* (Working Paper #139). Washington, DC: CALDER Center.
- Hansen, M., Backes, B., & Nelson, J. (2015). *Spilling over or watering down? Administrators' perceptions of Teach For America under a high-density placement strategy* (Unpublished Manuscript). Washington, DC: CALDER Center.
- Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., . . . Darwin, M. (2008). *Turning around chronically low-performing schools: A practice guide* (NCEE #2008-4020). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/wwc/publications/practiceguides>
- Hill, C., Bloom, H., Black, A., & Lipsey, M. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172–177.
- Hock, H., & Isenberg, E. (2012). *Methods for accounting for co-teaching in value-added models*. Washington, DC: Mathematica Policy Research.
- Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics, 1*(4), 85–108.

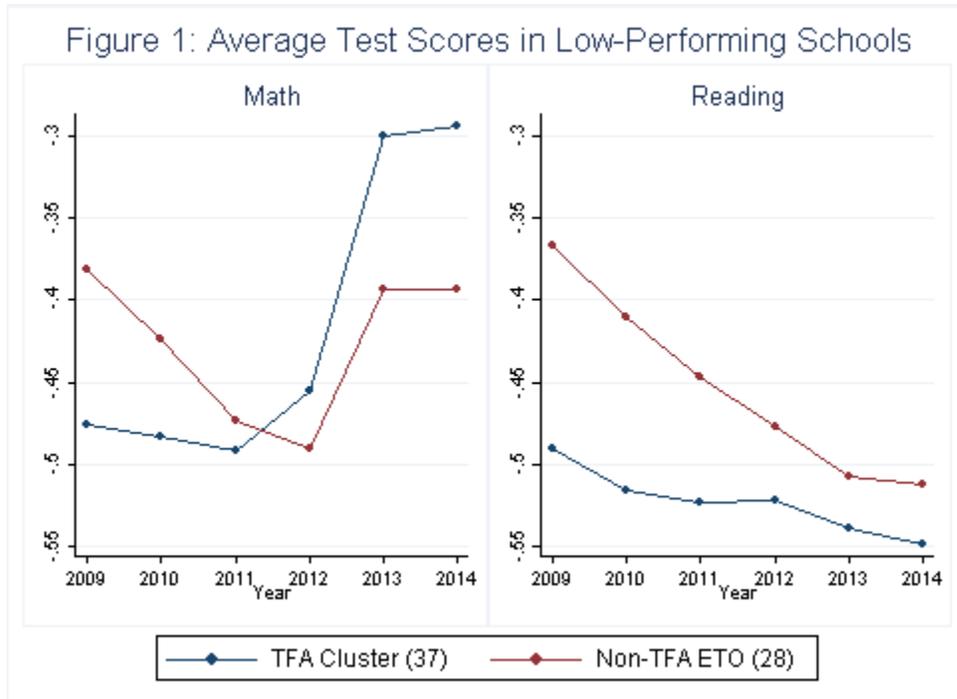
Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27(6), 615–631.

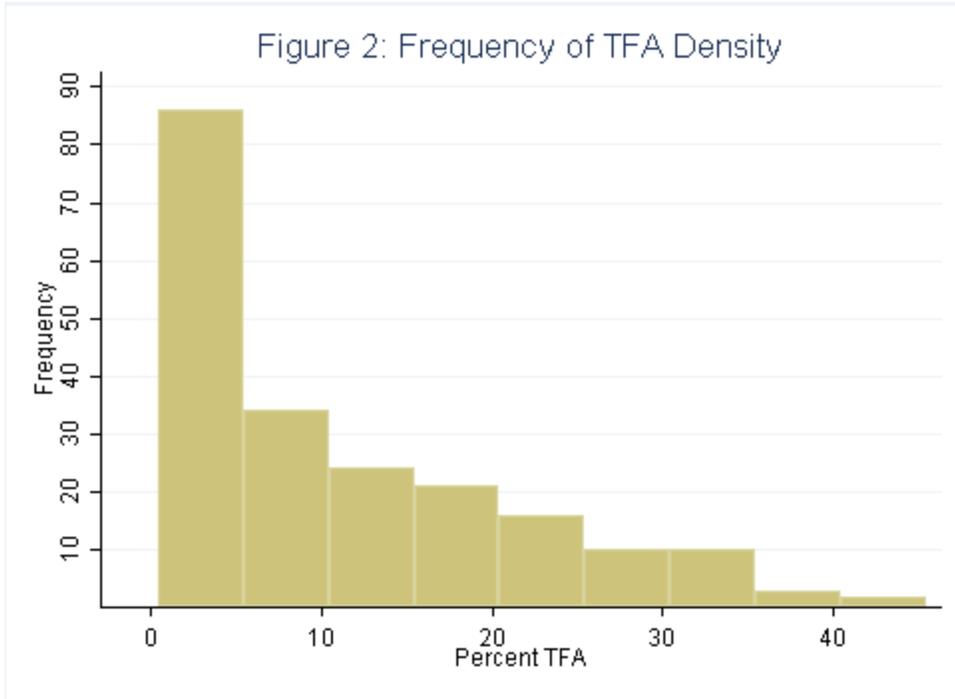
Koedel, C. (2009). An empirical analysis of teacher spillover effects in secondary school. *Economics of Education Review*, 28, 682–692.

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.

Yuan, K. (2015). A value-added study of teacher spillover effects across four core subjects in middle schools. *Education Policy Analysis Archives*, 23(38).

Figures





Note. Distribution of school-level percentages of TFA on staff for each school-year observation in the sample. School-year observations with zero TFA excluded.

Tables

Table 1. Active TFA Corps Member Assignments

	2008– 09	2009– 10	2010– 11	2011– 12	2012– 13	2013– 14
Total TFA Corps Members	88	88	137	222	271	290
Total Schools Containing Any TFA Corps Members	48	33	22	23	30	37
TFA as Proportion of School Teachers by School Type, Conditional on Containing TFA						
Elementary	3.43%	4.27%	9.33%	20.35%	13.77%	11.76%
Middle	3.98%	7.47%	8.53%	16.87%	16.87%	13.57%
High	1.73%	4.02%	13.64%	15.89%	14.88%	11.99%

Note. Proportions of schools teachers by school type are calculated among any schools containing any TFA corps members during that school year.

Table 2. Descriptive Statistics of the Analysis Samples

	Reading		Mathematics	
	Non-TFA Cluster School	TFA Cluster Schools	Non-TFA Cluster School	TFA Cluster Schools
Student-Level Variables				
Black	17.19%	68.51%	19.28%	70.31%
Hispanic	69.88%	29.51%	70.18%	28.05%
FRL Eligible	68.38%	87.07%	72.05%	88.16%
Mathematics Achievement			0.0431 (0.99)	-0.379 (0.99)
Reading Achievement	0.215 (0.96)	-0.322 (0.92)		
Unexcused Absences	4.31 (6.17)	7.86 (9.14)	4.39 (6.26)	8.04 (9.22)
Out-of-School Suspension Absences	0.39 (2.29)	1.35 (4.25)	0.43 (2.44)	1.53 (4.60)
Total Student-Year observations	972,421	107,264	868,372	92,313
Teacher-Level Variables				
MA Degree or Higher	36.27%	34.20%	33.62%	32.06%
Years of Experience	13.2 (9.84)	10.8 (9.29)	12.9 (9.84)	10.7 (9.27)
TFA Corps Member	0.13%	11.17%	0.15%	11.37%
Black	19.98%	51.51%	19.81%	49.74%
Hispanic	42.20%	17.50%	42.69%	19.79%
Total Teacher-Year Observations	27,860	3,178	21,998	2,461
Total Unique Schools	438	37	442	37

Note. TFA cluster schools are schools in which 2 or more new TFA corps members are placed in the same cohort for any cohort during or after the summer of 2009. *SDs* are reported in parentheses for outcome variables.

Table 3. Corps Members Assignments in the Analysis Sample

	2008–09	2009–10	2010–11	2011–12	2012–13	2013–14
Total TFA Corps Members	42	47	71	109	151	161
Total Schools Containing Any TFA Corps Members	30	22	20	20	24	33
TFA as Proportion of School Teachers by School Type, Conditional on Containing Any TFA						
Elementary	12.37%	13.24%	16.33%	25.58%	26.00%	22.41%
Middle	5.83%	10.46%	10.56%	21.47%	29.14%	19.68%
High	3.22%	7.39%	21.29%	23.36%	24.10%	21.27%
Average Classroom Characteristics for TFA Teachers						
Percent Black	70.24%	72.38%	78.21%	79.35%	77.44%	72.47%
Percent Hispanic	27.85%	26.34%	21.10%	19.60%	21.32%	26.10%
Percent FRL	86.34%	92.65%	92.05%	93.47%	92.14%	92.91%
Reading Achievement	-0.26 (0.84)	-0.52 (0.89)	-0.33 (0.83)	-0.31 (0.80)	-0.33 (0.83)	-0.29 (0.80)
Mathematics Achievement	-0.42 (0.90)	-0.44 (0.92)	-0.48 (0.96)	-0.28 (0.92)	-0.12 (0.93)	-0.22 (0.94)
Lagged Reading Achievement	- (0.84)	-0.50 (0.84)	-0.25 (0.81)	-0.22 (0.80)	-0.28 (0.82)	-0.23 (0.76)
Lagged Mathematics Achievement	- (0.89)	-0.43 (0.89)	-0.43 (0.88)	-0.32 (0.90)	-0.37 (0.87)	-0.41 (0.88)

Table 4. Baseline TFA Estimates

	Mathematics				Reading			
	1	2	3	4	5	6	7	8
TFA	0.102*** (0.020)	0.088*** (0.028)	0.109*** (0.018)	0.096*** (0.024)	0.015 (0.010)	0.014 (0.014)	0.022** (0.010)	0.027** (0.013)
TFA * 1 Year Experience		0.012 (0.036)		0.013 (0.031)		-0.010 (0.019)		-0.019 (0.018)
TFA * 2 Years Experience		0.052 (0.038)		0.038 (0.037)		0.006 (0.028)		-0.012 (0.026)
TFA * 3–4 Years Experience		0.051 (0.068)		0.063 (0.107)		0.051* (0.030)		0.036 (0.026)
1 Year Experience	0.046*** (0.017)	0.040 (0.026)	0.038** (0.016)	0.032 (0.023)	-0.001 (0.009)	0.002 (0.012)	-0.000 (0.009)	0.006 (0.011)
2 Years Experience	0.066*** (0.017)	0.056** (0.022)	0.080*** (0.018)	0.072*** (0.021)	0.025** (0.010)	0.024** (0.011)	0.014 (0.010)	0.017 (0.011)
3–4 Years Experience	0.059*** (0.017)	0.052** (0.020)	0.076*** (0.017)	0.070*** (0.020)	0.028*** (0.009)	0.027*** (0.010)	0.017* (0.009)	0.019* (0.010)
5–9 Years Experience	0.064*** (0.016)	0.058*** (0.019)	0.078*** (0.016)	0.072*** (0.019)	0.034*** (0.009)	0.034*** (0.010)	0.022** (0.009)	0.025** (0.010)
10+ Years Experience	0.045*** (0.016)	0.038** (0.019)	0.064*** (0.016)	0.058*** (0.019)	0.036*** (0.009)	0.035*** (0.010)	0.021** (0.009)	0.024** (0.010)
Observations	938,494	938,494	938,494	938,494	1,479,228	1,479,228	1,479,228	1,479,228
R-Squared	0.622	0.622	0.631	0.631	0.701	0.701	0.703	0.703
OLS	✓	✓			✓	✓		
School Fixed Effects			✓	✓			✓	✓

Note. *, **, *** = $p < 0.1, 0.05, 0.01$. School fixed effects models, with indicator variables on grade and year. Regression controls for student-level and class average demographics and cubic previous test scores, and their interactions with grade. Other controls include class size and teacher race and their interactions with grade.

Table 5. Spillover Effects and Student Outcomes

	Density: Entire School TFA Colleagues			
	Count	Percent	Count	Percent
	Mathematics		Reading	
	1	2	3	4
TFA	0.072*** (0.022)	0.102*** (0.028)	0.044*** (0.013)	0.045*** (0.016)
TFA Density	0.004** (0.002)	0.002 (0.001)	-0.000 (0.001)	-0.000 (0.001)
TFA * TFA Density	0.002* (0.001)	-0.001 (0.001)	-0.002*** (0.001)	-0.001 (0.001)
Observations	938,494	938,494	1,479,228	1,479,228
R-Squared	0.631	0.631	0.703	0.703

Note. *, **, *** = $p < 0.1, 0.05, 0.01$. School fixed effects models, with indicator variables on grade and year. Regression controls for student-level and class average demographics and cubic previous test scores, and their interactions with grade. Other controls include class size and teacher race and their interactions with grade. Percent TFA density measure is scaled as 0 to 100.

Table 6. Spillover Effects and Student Outcomes

	Density: TFA Grade Colleagues (in Elementary Grades) or TFA Subject Colleagues (in Middle / High Grades)			
	Count	Percent	Count	Percent
	Mathematics		Reading	
	1	2	3	4
TFA	0.091*** (0.025)	0.108*** (0.030)	0.029** (0.013)	0.030* (0.016)
TFA Density	0.008 (0.007)	0.001 (0.001)	-0.000 (0.002)	0.000 (0.000)
TFA * TFA Density	0.003 (0.006)	-0.001 (0.001)	-0.001 (0.002)	-0.000 (0.001)
Observations	938,494	938,494	1,479,228	1,479,228
R-Squared	0.631	0.631	0.703	0.703

Note. See note from Table 5.

Table 7. Changes in TFA Estimates Over Time

	TFA x Cohort Interactions		TFA x Year Interactions		
	Mathematics	Reading	Mathematics	Reading	
TFA	0.070** (0.029)	-0.009 (0.046)	TFA	0.090*** (0.028)	-0.005 (0.026)
TFA * 2007 Cohort	0.128 (0.093)	0.077 (0.051)	TFA * 2010 Year	(reference group)	
TFA * 2008 Cohort	(reference group)		TFA * 2011 Year	-0.058* (0.032)	0.026 (0.032)
TFA * 2009 Cohort	0.001 (0.050)	-0.005 (0.059)	TFA * 2012 Year	0.018 (0.042)	0.014 (0.033)
TFA * 2010 Cohort	0.009 (0.044)	0.024 (0.050)	TFA * 2013 Year	0.012 (0.037)	0.046* (0.028)
TFA * 2011 Cohort	-0.002 (0.056)	0.013 (0.051)	TFA * 2014 Year	0.053 (0.042)	0.030 (0.030)
TFA * 2012 Cohort	0.050 (0.056)	0.014 (0.050)	Observations	938,494	1,479,228
TFA * 2013 Cohort	0.075 (0.072)	0.092* (0.056)	R-Squared	0.631	0.703
Observations	938,494	1,479,228			
R-Squared	0.631	0.703			

Note. *, **, *** = $p < 0.1, 0.05, 0.01$. 2008 used as excluded cohort due to small sample size of 2007 cohort. *2007 Cohort* represents students taught by TFA corps members who were initially placed in the fall of 2007, and *2010 Year* indicates all students taught by any TFA corps member in the 2009–10 school year. Regressions are the same specification of columns 3 and 7 of Table 3.

Appendix: Data Cleaning Rules for Analysis

Various processes were undertaken during the course of the data analysis to create credible estimates of the TFA clustering effects. This appendix documents these various considerations and processes.

Tests Included in the Sample

Our final analysis sample spans Grades 4–10 and contains FCAT Reading, FCAT Mathematics, and Algebra EOC test scores. Each test score is standardized (z-scores) within year, grade, subject, and test type, relative to the district sample. Because pretest scores are needed as covariates in the regression, FCAT scores for third grade (the first tested grade) are used as pretest scores only (i.e., third grade observations do not appear in the analysis).

Up through the 2010–11 school year, it was possible for students in later grades to have taken two different mathematics tests in a year, the FCAT Mathematics and the Algebra EOC exam. For those students, we only use their FCAT Mathematics score. Starting in 2011–12, FCAT Mathematics is no longer tested in Grades 9 and 10, though the Algebra EOC exam continues to be tested at the conclusion of the algebra course (which some students may take for the first time as early as seventh grade or as late as 10th grade). In this case, Algebra EOC exam scores are only used when a student would otherwise be missing a mathematics test score in the current year (i.e., in Grades 9 and 10). For students taking algebra in seventh or eighth grade, the FCAT Mathematics score is used in those years, and those students' ninth-grade and 10th-grade mathematics observations are not included in the analysis.

Linking Students With Teachers

Course membership files in the data are used to identify the classes in which students receive instruction and the teachers to whom they are assigned. Students may be linked with multiple teachers in their course membership files (because of either switching classes midyear or taking multiple classes in the same subject, or due to coteaching arrangements).

Core Courses

When estimating value-added, we want to distribute student learning across all teachers in courses relevant to the tested subjects. As a result, it is important to distinguish between courses that focus on developing skills in tested subjects rather than elective courses that may only be tangentially related to a tested subject. For example, for mathematics value-added, we want to include an algebra course but exclude a computer science course that may be offered through the mathematics department and thus labeled in the data under a mathematics course code. We call courses focused on tested subjects *core courses* (CCs).

We developed the following two rules to help identify CCs for all students in the sample:

- (1) A course is flagged as a CC if 50% or more of the students in the district in that grade and year are enrolled in that same course (defined by the course code).
- (2) Any course that enrolls 10 or more students without being a CC (as determined by the first condition) is flagged as a CC for all students in that year and grade.

All non-CC student-teacher links are discarded. Teacher dosages (detailed below) are calculated based off of the remaining student-teacher links in CCs.

Estimating Regressions With Teacher Dosage

To properly attribute each teacher's contribution to a particular student's learning, we employ the Full Roster Method, developed by Hock and Isenberg (2012) of Mathematica Policy Research. This

method retains all student-teacher-course links labeled as CCs, and calculates a *teacher dosage* for each student-teacher link.

The M-DCPS data used for the analysis report course membership for students and teachers by terms, where each term represents half of the total exposure to a subject a student receives in a particular year (i.e., semesters). For each term, we distribute the term-subject dosage (0.5) across each of the student-teacher-course links observed. The term weights are added together to get the share of the total student-subject exposure that can be attributed to that student-teacher-course link such that the sum across all student-teacher-course links within a subject is 1. If a student leaves the sample at some point in the year, their student-subject exposure may be less than 1.

Consider the example presented in Appendix Table 1 below. Student A has four student-teacher-course links in English language arts for the 2011–12 school year. Three of these courses take place in the first term, the column labeled *# tchs in term 1* illustrates this value. Term 1's total student-subject exposure is 0.5, which is distributed across all three of these student-teacher-course links, the column labeled *Tch dos t1* represents the share of the term 1 dosage attributed to that student-teacher-course link. The same situation is true for term 2. Two of these courses are half-year courses and the other two are full-year courses; summing the dosage for each term gives more weight to the full-year courses and less weight to the half-year courses.

These full-year teacher dosages are incorporated into the value-added estimations as a student level analysis weight in Stata. Regressions are run using the *areg* command, which estimates dummy variables for each school fixed effect included in the model.

Appendix Table 1: Example of Assigning Teacher Dosages

Student	year	classid	tchid	Tch term1	Tch term2	# tchs in term1	# tchs in term2	Tch dose t1	Tch dose t2	tch_dosage
A	2012	843611	α	1	0	3		0.166667	0	0.166667
A	2012	843421	β	1	1	3	3	0.166667	0.166667	0.333333
A	2012	843495	β	1	1	3	3	0.166667	0.166667	0.333333
A	2012	843623	δ	0	1		3	0	0.166667	0.166667