Journal of Teacher Education
2019, Vol. 70(5) 453–471

DOI: 10.1177/0022487117752309
journals.sagepub.com/home/jte
SAGE

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

Ben Backes[1], Michael Hansen[2], Zeyu Xu[1], and Victoria Brady[3]

Abstract

This article examines Teach For America's (TFA) placement strategy in Miami-Dade County Public Schools, in which large numbers of TFA corps members are placed as clusters into a targeted set of disadvantaged schools, to investigate whether the large-scale infusion of TFA corps members into these schools induced broader improvements across the school. Using 6 years of administrative data from the district, we exploit variation in TFA density over time within schools to measure the extent to which increases in density were associated with improvements in student test scores. 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 direct effect of having larger numbers of classrooms staffed by effective TFA teachers. However, we do not find any evidence that the clustering strategy led to any spillovers on schoolwide performance.

Keywords

alternative certification, urban teacher education, school/teacher effectiveness, quantitative research

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 (Feeney, 2015). Many prior studies have evaluated the efficacy and turnover of TFA-placed teachers relative to other teachers in similar schools, though 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 (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Clark et al., 2013; Clark, Isenberg, Liu, & Makowski Zukiewicz, 2015; Glazerman, Mayer, & Decker, 2006; Kane, Rockoff, & Staiger, 2008; Xu, Hannaway, & Taylor, 2011).

Since 2009, TFA has used a novel 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 low-performing, disadvantaged schools. TFA adopted this new strategy in the region under the hypothesis that clustering corps members enables them to become more effective change agents in their schools (Lane, Lacefield-Parachini, & Isken, 2003) and could thus promote schoolwide improvement. This new strategy 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. Using administrative data on student test scores combined with TFA placement data, this article explores whether the quick, large-scale infusion of TFA corps members into these schools induced broader improvements across the school.

Prior Evidence on TFA

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's classroom performance, as measured by their ability to raise the test scores of their students, generally 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 et al., 2006) and quasi-experimental (Boyd et al., 2006; Kane et al., 2008; Xu et al., 2011) research designs. 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

[1]American Institutes for Research, Washington, DC, USA
[2]The Brookings Institution, Washington, DC, USA
[3]The New Teacher Project, Brooklyn, NY, USA

Corresponding Author:
Ben Backes, American Institutes for Research, 1050 Thomas Jefferson St., Washington, DC, USA.
Email: bbackes@air.org

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: Boyd et al. (2006) find first-year TFA middle school mathematics teachers outperform other beginning teachers by about 0.05 standard deviations of student achievement (*SD*); Clark et al. (2015) find lower elementary school students (Grades PK-2) scoring 0.12 *SD* higher in reading when taught by a TFA corps member.

The increased productivity of TFA teachers is presumed to be driven 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 with conflicting results. 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.), but 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.

Prior Evidence of Teacher Spillover

Based on our conversations with TFA regional personnel about the motivation behind their clustering strategy, they believed the presence of its corps members could be a catalyst for improvement schoolwide—not only in TFA-led classrooms—perhaps by transforming the school culture. Their hypothesis, though not drawn directly from the research literature on teacher education, is consistent with conceptualizations of new teachers as change agents promoting social justice in disadvantaged schools (e.g., Lane et al., 2003; McDonald & Zeichner, 2009) who could look to paired or clustered placements to provide professional support in taking on this task (similar to preservice placements, as in Bullough et al., 2002).

Based on these and other studies, one can reasonably hypothesize that TFA corps members may boost the performance of schools in which they are placed through the *spillover effects* on other teachers, thus extending their impacts beyond their own classrooms. In particular, there is a broad literature investigating the formation of ties and transmission of knowledge between teachers (e.g., Frank, Zhao, Penuel, Ellefson, & Porter, 2011; Penuel, Riel, Krause, & Frank, 2009; Spillane, Kim, & Frank, 2012) suggesting a scope through which effective TFA teachers may affect the performance of other teachers. In addition, Sun, Loeb, and Grissom (2017) find that when a teacher of above average effectiveness enters a school, the existing teachers in that teacher's grade experience a boost in performance.

Johnson (2015) recently reviewed the literature on teacher quality and makes the case for social capital between teachers (e.g., spillover) as an important mediator of teacher and school performance, arguing against the stereotypical "egg-crate" model of schooling commonly implied in this literature. In addition, spillover effects may become more pronounced as the number of TFA corps members clustered in a school increases: Jackson and Bruegmann (2009) find 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 gauge whether the presence of TFA teachers influences their peer teachers' performance. To do this, we leverage substantial variation in TFA density within targeted schools over time to measure the association between TFA density—our measure of the extent to which we may see spillover effects—and student achievement. 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. We therefore use within-school variation over time, generated by the clustering strategy, to measure spillover effects for this study.

School Turnaround Through Human Resource Turnover

The theory of action linking infusions of relatively effective teachers with school improvement is well established in the literature on school turnaround and has been heavily influenced by ideas of business turnaround in organizational management research (see Rhim, Kowal, Hassel, & Hassel, 2007). According to this literature, low-performing organizations can improve through the use of selective turnover, removing individuals who are unwilling or unable to improve, to better promote a performance-oriented culture in the organization (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010; Collins, 2001; Kearns, 2010). These ideas have been ingrained in models of school improvement for years. One of the four federally prescribed models of school turnaround under the 2009 Race to the Top initiative (the *turnaround* model) explicitly calls for at least 50% turnover among the low-performing school's teaching staff. Before that, one of the models of school restructuring under the No Child Left Behind act similarly called for *Turnarounds with New Leaders and Staff*.

Despite its policy endorsement, this strategy is based on a relatively weak evidence base. The Institute of Education Science's *Practice Guide* on the topic (Herman et al., 2008) reviewed both relevant business literature and the school improvement literature and, while suggestive, found no convincing evidence of a causal relationship between selective turnover and school improvement. 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. Hansen (2013) investigates quick improvements in low-performing schools, decomposing contributions from teachers new to the school versus improvements among stable staff and finds evidence that both were important contributors to achievement gains. Though the author does not investigate the mechanisms through which stable teachers improved, performance spillovers from new, highly effective colleagues onto previously less effective teachers could be one mechanism that is consistent with the observed patterns of improvement among teachers.

This approach to school turnaround is not only useful context for the theory behind TFA's clustering strategy but it is also an integral part of how the strategy was implemented in Miami. 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. This partnership between TFA and the ETO (described in more detail in the following section) 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*, which required replacing at least 50% of teachers in the low-performing school.

This Study's Contributions

This study investigates the role of TFA spillover on both non-TFA and other TFA colleagues. Our primary research question is whether increases in the TFA share of a school's teaching staff are associated with improvements in the effectiveness of either TFA or non-TFA teachers, as measured by the performance of their students. We use administrative data from M-DCPS for the six school years between 2008-2009 and 2013-2014, with school fixed effects to leverage the variation in TFA density over time within schools.[1] In 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 that the performance of TFA corps members in their own classrooms contributed to a modest portion of these gains. However, we do not find any evidence that any of these gains can be explained by the share of TFA corps members in a school.

TFA Placement in Miami-Dade

TFA started placing corps members in M-DCPS in 2003, with 35 initial placements.[2] 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.

Motivations for TFA Clustering

Beginning with the 2009-2010 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 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.[3] 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. 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.

A counterhypothesis suggesting that clustering may inadvertently cause harm to targeted schools is also possible, though based on our conversations with TFA regional staff and district personnel, this was not considered when designing and implementing the clustering strategy. Low-performing schools may have a hard time supporting the large influx of relatively short-term, novice teachers. Alternatively, perhaps part of the TFA effect documented in prior studies could be due to coaching from senior teachers in their schools, and if

Table 1. TFA School Assignments.

	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014
Total TFA corps members	88	88	137	222	271	290
Total schools containing any TFA corps members	48	33	22	23	30	37
Mean total TFA (corps + alumni) per TFA placement school	1.9 (1.2)	2.8 (2.6)	6.8 (5.8)	10.3 (7.1)	9.9 (9.5)	9.4 (8.5)
TFA as percent of school teachers by school type, conditional on containing TFA						
Elementary	3.4%	4.3%	9.3%	20.4%	13.8%	11.8%
Middle	4.0%	7.5%	8.5%	16.9%	16.9%	13.6%
High	1.7%	4.0%	13.6%	15.9%	14.9%	12.0%

Note. Proportions of schools teachers by school type are calculated among any schools containing any TFA corps members during that school year. Standard deviations in parenthesis. TFA = Teach For America.

cluster schools have fewer senior teachers and more TFA teachers in need of coaching, TFA teachers' performance in cluster schools could suffer. Whether TFA performance shows any relationship with the placement of its corps members is an empirical question that we examine here.

TFA Partners With the ETO

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. The numbers of schools targeted by the each program and their overlap demonstrate the common objectives between TFA's clustering strategy and the ETO: 31 of the 37 TFA cluster schools in our sample are also ETO schools. In addition, there are 28 ETO schools that are not TFA cluster schools. Both TFA and the ETO entities viewed this as a mutually beneficial partnership—The ETO valued the flow of corps members to vacancies that are otherwise difficult to staff,[4] while TFA viewed this as a way to strategically target their efforts in the highest need schools, which was necessary to maximize their impact on the highest need students. This partnership further accelerated the growth of the total number of corps members working in the district.

Schools under the guidance of ETO received funding through the U.S. Department of Education's School Improvement Grant (SIG) Fund. Placement of TFA corps members in some schools was one part of a multipronged approach to school improvement in the district. According to the ETO, examples of interventions that SIG funds supported were incentive pay to aid with the retention of effective administrators and teachers, extending the learning day, receiving on-site professional development, and hiring TFA corps members.[5] Based on conversations with district personnel in the ETO, TFA hiring was one of several interventions that the ETO used in turnaround schools, and in every school multiple interventions were carried out simultaneously (i.e., TFA hiring was not the only turnaround strategy in any school). Because we find evidence that ETO identification was associated with later gains in math scores, as discussed below, we explicitly control for ETO identification to avoid conflating ETO-driven effects with TFA clustering.

Increased TFA Density in Targeted Schools

The growth of the TFA corps and its density are readily apparent in the placement numbers during the six school years of the data used for this analysis. Table 1 presents TFA corps member assignment figures over time. In the 2008-2009 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 about 10 TFA teachers per school where there was any presence.[6] The net result was a jump in the proportion of TFA in placement schools, going from 2% to 4% in 2008-2009 to 12% to 14% in 2013-2014.

Conceptualizing Spillover Effects

Theory of Action

The theory behind spillover presumes that all teachers affect and learn from each other. Sun et al. (2017) frame the possible mechanisms for spillover between teachers into knowledge transfer and social pressure. In empirical data, multiple studies have shown that a teacher's effectiveness depends in part on how effective her colleagues are and are associated with changes in peer composition (e.g., Jackson & Bruegmann, 2009; Koedel, 2009; Sun et al., 2017). The specific mechanisms driving such learning may include informal learning between teachers (Eraut, 2004; Frank et al., 2011) and formal mentoring arrangements promoting professional development (Papay, Taylor, Tyler, & Laski, 2016), and may be moderated by a supportive professional environment (Kraft & Papay, 2014) that leads to skill development. Other mechanisms have also been speculated as important transmitters across teachers, which could include the sharing of instructional resources, coaching each other, or simply motivating each other (Johnson, 2015).

As we show in the Results section, TFA teachers are more effective than the average non-TFA teacher in M-DCPS in math and reading in our sample, meaning non-TFA teachers learning from TFA and becoming more effective could be plausible. Because TFA teachers are more effective than non-TFA in the sample, Sun et al.'s (2017) finding that the introduction of peers more effective than the current teachers in a school raises the effectiveness of the existing teachers is especially relevant. The infusion of TFA clusters could feasibly promote greater student learning through allowing more frequent interactions and learning opportunities from these relatively effective teachers. During in-person interviews, multiple school principals discussed ways in which it was plausible that TFA teachers could influence their colleagues' practice. The most common opportunity principals cited was 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. Several principals participating in the interviews also indicated specific mechanisms through which they felt TFA teachers positively influenced the school's professional culture, including high energy levels, high expectations, and outreach to parents, which they cited as being noticed and modeled by other teachers. Thus, if these behavioral spillovers translate into student achievement, a TFA spillover effect among all teachers broadly may be plausible.[7]

One factor that could limit the potential spillover from TFA teachers to non-TFA teachers is a difference in background between the two groups. The typical TFA teacher is much more likely to be young, white, and male than the typical non-TFA teacher in the same schools (Hansen, Backes, & Brady, 2016). In addition, the interviews conducted by Trujillo and Scott (2014) suggest that TFA corps members may have views that strongly diverge from traditionally trained teachers on topics like the strength of teachers unions and teacher accountability. Non-TFA teachers are also aware that TFA teachers are unlikely to remain at the same school for multiple years, especially those with difficult teaching assignments such as multigrade assignments (Donaldson & Johnson, 2010, 2011) as well as younger TFA teachers (Donaldson, 2012). Spillane et al. (2012) find that teachers of the same race, gender, and career stage (as measured by teacher experience) are more likely to give and receive advice to each other than to those who are observationally dissimilar. Consistent with Spillane et al. (2012), Jackson and Bruegmann (2009) 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. If these findings hold for influxes of TFA corps members, where almost all of these corps members are novices, then the density of TFA may have the strongest effect on other TFA corps members, and only a moderate or null 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 either TFA or non-TFA teachers.

Modeling Spillover Effects

The spillover effects of interest for this article 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: (a) Through what group are spillover effects transmitted? (b) How is the concentration of TFA teachers measured in the group? Both of these dimensions are discussed in turn below.

First, how are spillover effects transmitted? The clustering strategy, as implemented in M-DCPS, focused on increasing the presence of TFA teachers 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 the TFA 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. At the elementary school level, as discussed above, Spillane et al. (2012) have found that teaching in the same grade is predictive of the formation of new ties between teachers. In these cases, we would expect that a parameterization that defines the peer group more narrowly (at the grade or subject level) 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, and 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 concentration 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 as our preferred specification because it most closely mirrors the weighted average effectiveness measure that Jackson and Bruegmann (2009) use.

Finally, we note that this study looks for evidence of fairly large spillovers associated with the concentration of TFA. The estimated magnitude of productivity spillovers across teachers is relatively small in the Jackson and Bruegmann (2009) study. They estimate that an increase of 1 *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 under the assumption that the spillovers estimated in Jackson and Bruegmann (2009) generalize to M-DCPS 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, 2009, consider other teachers within a grade and school as peers in their sample of elementary school students). Thus, for a grade with five teachers (plugging $n = 5$ into the equation above), the expected increase in other students' test scores due to replacing one with a TFA teacher would be 0.003 *SD*, an effect that would be too small to detect given the standard errors in our estimates. While we do not have any direct evidence on whether to expect the Jackson and Bruegmann estimate to generalize to our setting, even if the true spillover effects were twice as large in M-DCPS, detecting spillover effects due to TFA would be challenging.

The minimum detectable effect for a hypothetical 20 percentage point increase in the share of TFA teachers (the effect of, for example, replacing one out of five teachers with a TFA teacher) in our data is many times larger, at 0.04 *SD* (obtained by multiplying the standard error in Table 5, column 2 by 1.96 × 20). Though we acknowledge this magnitude represents a large spillover effect, it remains important to investigate the effectiveness of a strategy of improving low-performing schools with clusters of TFA. The hypothesis that TFA teachers (or, generalizing our analysis, other relatively effective new teachers as in Lane et al., 2003) act as a catalyst for broader improvement among all teachers in a school implicitly supposes that modest changes in staffing could lead to increased productivity in many classrooms. Indeed, as discussed above, this was one of the original stated motivations from TFA about why they decided to implement the clustering strategy in M-DCPS. If spillover is to play an influential role in a low-performing school's turnaround strategy (a plausible explanation for the evidence in Author, 2013), it must be quite large and detectable. Though our study could miss smaller spillover effects that may possibly be present due to data limitations, smaller effects would arguably be inconsequential in affecting schoolwide change of the magnitude desired for low-performing schools. Moreover, the absence of detectable effects would weaken the justification for clustering together corps members under this strategy.

Data

We use detailed student-level administrative data that cover M-DCPS students linked to their teachers for six school years (2008-2009 through 2013-2014). 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).[8] Students' FCAT scale scores are converted to *z*-scores based on the mean and *SD* for that particular year-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.[9]

Teacher personnel files in the M-DCPS data contain information on teachers' experience levels, education attainment, demographics, and other supplemental background variables. These are likewise used as covariates for various models in the analysis that follows. One variable included in the data is whether the teacher is affiliated with TFA. Given the importance of this variable in the analysis, we externally

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	−0.379
			(0.99)	(0.99)
Reading achievement	0.215	−0.322		
	(0.96)	(0.92)		
Unexcused absences	4.31	7.86	4.39	8.04
	(6.17)	(9.14)	(6.26)	(9.22)
Out-of-school suspension absences	0.39	1.35	0.43	1.53
	(2.29)	(4.25)	(2.44)	(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	10.8	12.9	10.7
	(9.84)	(9.29)	(9.84)	(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 two 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. TFA = Teach For America; FRL = free or reduced-price lunch.

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). One weakness of the data is the reliance on student outcomes. Thus, we cannot observe integration between TFA and non-TFA teachers, but rather measure whether teaching in the same school as productive TFA teachers increases other teachers' ability to raise the test scores of their students.

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. Active corps members comprise about 88% of the TFA observations in the analysis sample (the remaining 12% of TFA observations are from years in which TFA teachers are considered alumni). We flag observations of alumni as TFA because all were active corps members in the district at some point and generally continue some form of engagement with TFA as alumni, although the decision to flag them does not substantively affect our results.

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. In contrast to the noncluster schools in Miami, where Hispanic students constitute a majority, more than three fourths of students in cluster schools are Black. 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.

Table 3. TFA Assignments in the Analysis Sample.

	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014
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	−0.52	−0.33	−0.31	−0.33	−0.29
	(0.84)	(0.89)	(0.83)	(0.80)	(0.83)	(0.80)
Mathematics achievement	−0.42	−0.44	−0.48	−0.28	−0.12	−0.22
	(0.90)	(0.92)	(0.96)	(0.92)	(0.93)	(0.94)
Lagged reading achievement	—	−0.50	−0.25	−0.22	−0.28	−0.23
	—	(0.84)	(0.81)	(0.80)	(0.82)	(0.76)
Lagged mathematics achievement	—	−0.43	−0.43	−0.32	−0.37	−0.41
	—	(0.89)	(0.88)	(0.90)	(0.87)	(0.88)

Note. TFA = Teach For America; FRL = free or reduced-price lunch.

By construction, the share of TFA teachers is much higher in the cluster sample.[10] 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). These figures indicate TFA being overrepresented among tested grades and subjects, which is unsurprising because they 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.

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 teachers. The posttest scores jump considerably in the last 3 years of data, an increase of well over 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 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 2008-2009 performance. Note that most, but not all, of the 37 TFA cluster schools are also considered ETO schools,

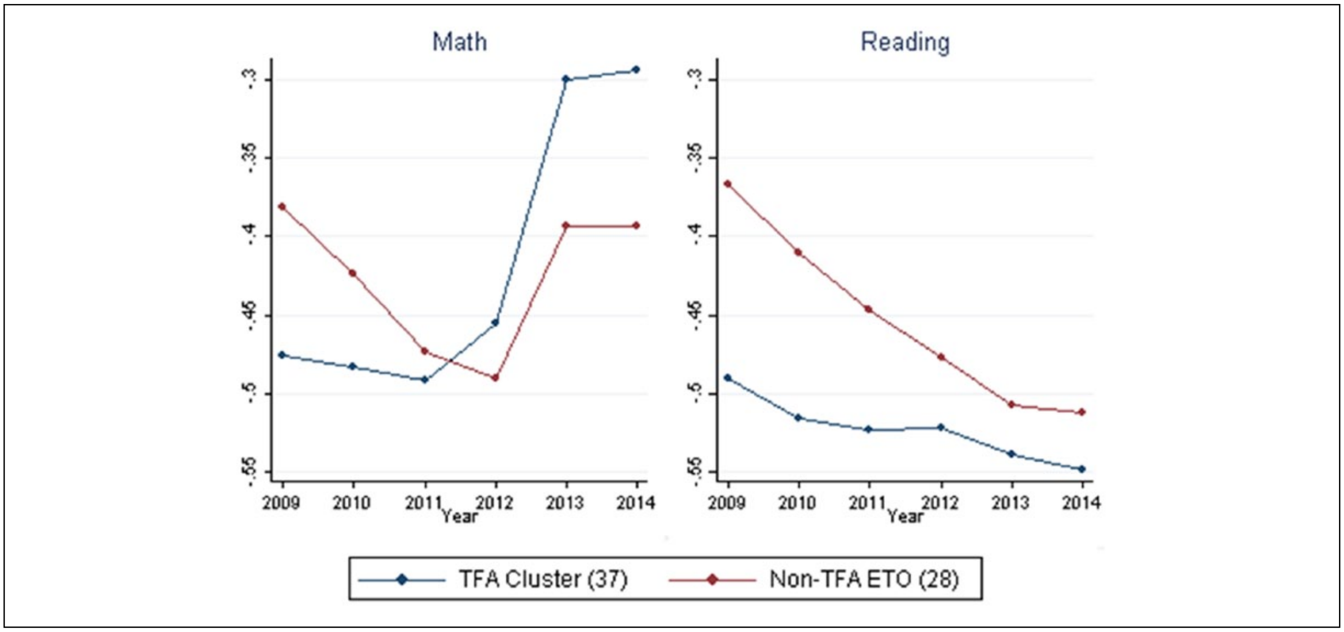

Figure 1. Average test scores in low-performing schools.
Note. TFA = Teach For America; ETO = Education Transformation Office.

whereas the noncluster ETO group contains 28 schools.[11] Hence, to avoid attributing this rise in achievement solely to the TFA clustering strategy, we control for year-specific ETO effects in our analysis.[12]

Empirical Strategy

To model the relationship between TFA's clustering strategy and 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_{i}), classroom characteristics (X_{ct}), and a school fixed effect (γ_s). Studies of TFA effectiveness generally estimate an equation similar to the following:[13]

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

with the vector of prior-year test scores A_{it-1} containing cubic functions of prior test scores in both reading and mathematics; the vector of student characteristics X_{it} including race, gender, FRL eligibility, LEP status, and mental, physical, or emotional disability status; the vector of classroom characteristics X_{ct} including class size, classroom-level averages of prior-year test scores, classroom-level averages of each of the student characteristics listed above; and characteristics of the classroom teacher including teacher race, 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, and the estimating equation additionally includes indicator variables for grades and years. The coefficient β_3 then estimates the expected change in test scores associated with being in a TFA classroom relative to non-TFA classrooms with similar student backgrounds and test scores.

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:

$$\begin{aligned} A_{ist} &= \alpha A_{it-1} + \beta_1 X_{it} + \beta_2 X_{ct} + \gamma_s + \beta_3 TFA_{ct} \\ &\quad + \beta_4 TFA\_DENSITY_{st} + \beta_5 TFA\_DENSITY_{st} \\ &\quad \times TFA_{it} + \beta_6 ETO_s \times YEAR_t + \varepsilon_{ist}. \end{aligned} \quad (2)$$

In the equation above, *TFA_DENSITY* is a measure of TFA density among the relevant peer group (referenced in separate regressions against the whole school versus colleagues in the same grade or subject as described previously). Changes in student learning associated with increases in density constitute our test for the existence of spillover. What we capture when estimating Equation 2 is whether an individual with given prior test scores and demographic information is estimated to perform better when located in a TFA-dense setting than otherwise. Thus, there are three coefficients of interest. First, β_3, which measures the direct effect of TFA teachers in their own classrooms. If β_3 is positive, then overall school performance will increase as the number of TFA increases due to this direct effect alone. The focus of the article is whether there is any *additional* schoolwide change in test scores associated with increases in the concentration

Table 4. Baseline TFA Estimates.

	Math		Reading	
	1	2	3	4
TFA	0.10***	0.11***	0.02	0.02**
	(0.02)	(0.02)	(0.01)	(0.01)
1 year experience	0.05***	0.04**	−0.00	−0.00
	(0.02)	(0.02)	(0.01)	(0.01)
2 years experience	0.07***	0.08***	0.03**	0.01
	(0.02)	(0.02)	(0.01)	(0.01)
3-4 years experience	0.06***	0.08***	0.03***	0.02*
	(0.02)	(0.02)	(0.01)	(0.01)
5-9 years experience	0.06***	0.08***	0.03***	0.02**
	(0.02)	(0.02)	(0.01)	(0.01)
10+ years experience	0.05***	0.06***	0.04***	0.02**
	(0.02)	(0.02)	(0.01)	(0.01)
Observations	938,494	938,494	1,479,228	1,479,228
R^2	.62	.63	.70	.70
OLS	✓		✓	
School fixed effects		✓		✓

Note. 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. TFA = Teach For America; OLS = ordinary least squares.
**p* < .1. ***p* < .05. ****p* <. 01.

of TFA, over and above TFA impacts in their own classrooms. These are estimated by β_4 and β_5, with β_4 capturing the change in estimated performance in non-TFA classrooms and β_5 in TFA classrooms.[14] As described above, we also include ETO School × Year interactions to avoid attributing increases in school performance due to ETO interventions to TFA density.[15] All standard errors are clustered at the school level.

Equation 2 is a variation of a difference-in-difference (DD) design, with the inclusion of school fixed effects meaning that the variation in schools' densities of TFA teachers over time is driving the resulting estimates. 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.[16] Using a continuous measure of density allows us to more precisely describe the prevalence of TFA at a given school.

Equation 2 would have a causal interpretation if schoolwide performance in treated schools would have evolved in a similar manner as control schools in the absence of treatment. This assumption may be unlikely to hold due to nonrandom selection into treatment as well as concurrent interventions in the district led by ETO. For example, if principals who are more favorable to TFA choose to participate, then the coefficient on TFA density may overstate what the effect would be for a randomly selected school. While the presence of school fixed effects in the model should capture any unobservable time-invariant differences across schools, changes in schools' unobservable likelihood to hire TFA teachers over time (which could be possible, for example, due to principal turnover during the study period) could leave some residual confounding bias in our estimates. Hence, we cannot claim the spillover estimates presented below have a causal relationship with student achievement.

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. The first column shows a basic ordinary least squares (OLS) regression while controlling for TFA and teacher experience. 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 column 3.[17] Columns 2 (mathematics) and 4 (reading) add school fixed effects, which are frequently used in studies of TFA to account for fixed differences across schools (see, for example, Boyd et al., 2006, and discussion within) to ensure that the TFA coefficient is not being downwardly biased in OLS when corps members are placed in relatively disadvantaged schools. However, in Table 4 we generally find similar results across the OLS and school fixed effects estimates.

Table 5. Spillover Effects and Student Outcomes: School Level.

	Density: Entire school TFA colleagues			
	Count	%	Count	%
	Mathematics		Reading	
	1	2	3	4
TFA	0.072***	0.102***	0.044***	0.045***
	(0.022)	(0.028)	(0.013)	(0.016)
TFA density	0.004**	0.002	−0.000	−0.000
	(0.002)	(0.001)	(0.001)	(0.001)
TFA × TFA	0.002*	−0.001	−0.002***	−0.001
Density	(0.001)	(0.001)	(0.001)	(0.001)
Observations	938,494	938,494	1,479,228	1,479,228
R^2	.63	.63	.70	.70

Note. 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 as well as interactions between year and ETO status. Percent TFA density measure is scaled as 0 to 100. TFA = Teach For America.
*p < .1. **p < .05. ***p <. 01.

In the school fixed effects specification, TFA corps members now significantly outperform non-TFA colleagues in reading by a modest 0.02 *SD*. The magnitude of our TFA mathematics effect—about 11% of a *SD* of student learning on standardized test scores—falls roughly in the middle of previous studies. Relative to the articles discussed earlier, our estimate is somewhat smaller than Glazerman et al. (2006) (0.15 *SD*) and Xu et al. (2011) (0.13 *SD*), and larger than Clark et al. (0.07 *SD*), Kane et al. (2008) (0.02 *SD*), and Boyd et al. (2006) (no effect).

For ease of interpretation, we convert this point estimate for math into months of learning. 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, Bloom, Black, and Lipsey (2008) which uses nationally normed standardized test results.[18] 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. In addition, we make this conversion to months of learning using average gains observed in the FCAT data based on technical reports for all test-takers in the state.[19] Using these estimates to make this conversion slightly tempers the TFA effect, to 2.6 months of learning (a 26% boost in learning relative to a 10-month calendar).

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 0.02 *SD*, comparing this against the magnitudes of the estimates from greater experience (in column 4) shows the TFA effect to be roughly equal in magnitude to the effect of having much more senior teachers in the classroom in these 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 article: whether the density of TFA members in a school is associated with 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.[20] 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 with the other point estimates of TFA on mathematics in Tables 4 to 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 find no evidence supporting the "critical mass" hypothesis for cultural transformation.[21]

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 spillovers. Those specifications that are significantly different are from the less-preferred count specifications. Our critical mass threshold explorations turned up no

Table 6. Spillover Effects and Student Outcomes: Grade/Subject Level.

	Density: TFA grade colleagues (in elementary grades) or TFA subject colleagues (in middle/high grades)			
	Count	%	Count	%
	Mathematics		Reading	
	1	2	3	4
TFA	0.091***	0.108***	0.029**	0.030*
	(0.025)	(0.030)	(0.013)	(0.016)
TFA density	0.008	0.001	−0.000	0.000
	(0.007)	(0.001)	(0.002)	(0.000)
TFA × TFA Density	0.003	−0.001	−0.001	−0.000
	(0.006)	(0.001)	(0.002)	(0.001)
Observations	938,494	938,494	1,479,228	1,479,228
R^2	.63	.63	.70	.70

Note. See note from Table 5. TFA = Teach For America.

significant differences on these interacted variables either. We conclude that there is no evidence that the performance of TFA teachers are related to the concentration of TFA in their schools.

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. One potential concern about a rapid expansion in TFA placements over time is that if there are a limited number of high-quality TFA applicants, then quickly increasing the number of corps members could lead to a dilution in quality. If the clustering strategy led to a concurrent decrease in quality, our estimates would then understate spillover because the empirical model estimates a single TFA effect over the entire period. Not only is this possible drop in quality a theoretical 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 district administrators interviewed.[22]

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.[23] 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.[24]

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 relationship between TFA concentration and teacher effectiveness in the data. Based on Table 7, we have no reason to believe a concurrent decrease in quality deflates spillover that may be otherwise present. One could consider it a disappointment that we fail to find evidence that the clustering strategy was responsible for noticeable achievement gains. This could be due to a number of reasons, such as the limited size of spillover in practice (e.g., Jackson & Bruegmann, 2009) or lack of integration between TFA and non-TFA teachers in high-density schools (a common concern voiced among principals during interviews). However, there may be some other mitigating factors behind the interpretation of these results; we briefly discuss these here.

The first is related to the magnitude of the effect that we can detect given the precision of our estimates. As described previously, the spillover estimates in the Jackson and Bruegmann (2009) study are very small, where our minimum detectable effects are many times larger. In light of this

Table 7. Changes in TFA Estimates Over Time.

	Mathematics	Reading
TFA × Cohort interactions		
TFA	0.07**	−0.01
	(0.03)	(0.05)
TFA × 2007 Cohort	0.13	0.08
	(0.09)	(0.05)
TFA × 2008 Cohort	(reference group)	
TFA × 2009 Cohort	0.00	−0.01
	(0.05)	(0.06)
TFA × 2010 Cohort	0.01	0.02
	(0.04)	(0.05)
TFA × 2011 Cohort	−0.00	0.01
	(0.06)	(0.05)
TFA × 2012 Cohort	0.05	0.01
	(0.06)	(0.05)
TFA × 2013 Cohort	0.08	0.09*
	(0.07)	(0.06)
Observations	938,494	1,479,228
R^2	.63	.70
TFA × Year interactions		
TFA	0.09***	−0.0
	(0.03)	(0.03)
TFA × 2010 Year	(reference group)	
TFA × 2011 Year	−0.06*	0.03
	(0.03)	(0.03)
TFA × 2012 Year	0.02	0.01
	(0.04)	(0.033)
TFA × 2013 Year	0.01	0.05*
	(0.04)	(0.03)
TFA × 2014 Year	0.05	0.03
	(0.04)	(0.03)
Observations	938,494	1,479,228
R^2	.63	.70

Note. 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 2009-2010 school year. Regressions are the same specification of columns 2 and 4 of Table 4. TFA = Teach For America.
**p* < .1. ***p* < .05. ****p* <. 01.

minimum detectable effect, it is no surprise that we cannot detect spillovers 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 with a novice (based on the point estimate in column 2 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 associated with 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. Spillovers 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. The use of administrative data in this article does not capture any of the ways that teachers interact with each other, so to the extent that there are spillovers between teachers that are not reflected by student test scores, we would not be able to measure them. Thus, we are unable to investigate the mechanisms driving our results. For example, it could be the case that TFA teachers do not become integrated with other teachers in schools with large TFA populations, but our data do not speak to that.

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 et al., 2016), as well as and TFA impacts on nontested student outcomes (Backes & Hansen, in press). Summarizing across these studies, it appears that the primary impact of expanding TFA's presence in Miami was through the direct effect of an increased share of classrooms being taught by TFA teachers. Other hypothesized outcomes, such as improved retention of TFA corps members, failed to materialize.

Conclusion

The research question motivating this study asks whether increases in TFA corps members are associated with changes in schoolwide performance outside their classrooms 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 this question. 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 a meaningful relationship between the density of TFA in a school and the performance of other teachers in the school as measured by student test scores—neither for non-TFA teachers nor on TFA corps members in the same schools. We explore a variety of specifications of the TFA density measure, and virtually all result in no significant differences associated with these changes induced by clustering. However, we do find robust evidence of TFA effects on mathematics test scores exceeding 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.[25]

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: (a) to accelerate TFA's influence in student outcomes in particularly disadvantaged settings and (b) 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. Because we find that TFA teachers are more effective than the average non-TFA teacher in M-DCPS, TFA's increased presence has made a difference in student learning in the district. For example, if 20% to 30% of math teachers in a school are TFA (Table 3), and the average TFA impact is 11% of a standard deviation, then TFA schools should perform about 2% to 3% of a standard deviation above a non-TFA school. This represents 20% to 30% of the gap between TFA and non-TFA ETO schools in the recent years in Figure 1, with the remainder of increased performance in TFA schools being unexplained.

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 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.

It is also notable that there is no evidence to suggest that TFA effectiveness in high-density schools was watered down at all. The district's ETO commonly used TFA in part to fill hard-to-staff positions in schools implementing the *turnaround model* that required high levels of staff turnover. Though some could consider relying on so many inexperienced teachers a real risk for students in these schools, we do not find any evidence suggesting that this strategy harmed students. Rather, our results show TFA teachers outperformed comparison teachers in their schools regardless of the placement strategy; thus, ETO's solution of TFA placements appears to have been effective both in implementing the strategy and in delivering stronger student achievement to students in TFA teachers' classrooms.

Finally, these findings have broader implications beyond TFA as well; we briefly discuss two here. First, education scholars have commonly promoted a view of new teachers—particularly in disadvantaged or urban settings—as change agents in the schools they teach, both for the students they serve (Catapano, 2006; McDonald & Zeichner, 2009) and among their teacher colleagues (Lane et al., 2003). The evidence presented here suggests new teachers likely have a relatively modest scope of influencing change in their new schools. Even where clusters of relatively effective novice teachers are placed jointly (thus, presumably increasing their potential to take risks and influence others' practice, as in Bullough et al., 2002), looking to new teachers as the catalysts for schoolwide improvement likely expects too much from individuals in a larger institutional context.

Second, states and districts continue to cope with how best to bring about improvements in chronically low-performing schools, and human capital strategies to build a committed staff are common. Some strategies include recruiting highly effective teachers (or in some cases, teams of them) to teach in low-performing schools in part to help establish new professional norms (e.g., Glazerman, Protik, Teh, Bruch, & Max, 2013). Generalizing our findings to other highly effective teachers placed in low-performing schools, it is unlikely that such placements alone would generate schoolwide improvements through spillover (though this does not rule out the possibility of such transfers having large impact on students in their own classrooms).

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 to 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 2010-2011 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 2011-2012, 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).

Following implementation steps described in Hock and Isenberg (2017), 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 (2017) 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. Student A has four student–teacher–course links in English language arts for the 2011-2012 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.

Table A1. 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	0.1666667
A	2012	843421	β	1	1	3	3	0.166667	0.1666667	0.3333333
A	2012	843495	β	1	1	3	3	0.166667	0.1666667	0.3333333
A	2012	843623	δ	0	1		3	0	0.1666667	0.1666667

Acknowledgments

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.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was made possible in part by generous financial support from 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. 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.

Notes

1. A component of this study's larger project included interviews with 12 school and more than five district personnel in M-DCPS who had worked closely with TFA during the study period. Though we occasionally reference these conversations to add to the narration of this article's inquiry, a full treatment of our findings from these interviews is beyond the scope of this article; we refer the reader to Hansen, Backes, and Nelson (2015) for a full report of these interviews.
2. 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.
3. 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, for example, Hansen, Backes, & Brady, 2016).
4. TFA placements under the clustering strategy are still heavily dependent on 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).
5. A full analysis of turnaround interventions in the district is beyond the scope of this article. For additional information on school turnaround in the district, please see SIG documentation from the district (ttp://www.aspendrl.org/portal/browse/DocumentDetail?documentId=866&download) and the ETO website (http://eto.dadeschools.net/).
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 3 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.
7. 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.
8. From the 2008-2009 school year through the 2010-2011 school year, all students Grades 3 to 10 took the FCAT in both mathematics and reading. However, with the introduction of

End-of-Course (EOC) exams in 2011-2012, the mathematics portion of the FCAT will only be administered to Grades 3 to 8 from 2011-2012 forward. For students taking an EOC exam in 2011-2012 through 2013-2014 (e.g., Algebra I), we consider their previous year's FCAT score to be their lagged test score. See Appendix for more information.

9. 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.
10. 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.
11. The list of schools considered ETO by the district has grown over the last several years; we identify a school as 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.
12. 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.
13. Studies using this approach include Boyd, Grossman, Lankford, Loeb, and Wyckoff (2006); Clark et al. (2013); Glazerman, Mayer, and Decker (2006); and Kane, Rockoff, and Staiger (2008).
14. To allow for students with multiple teachers, regressions are run using the Full Roster Method (Hock & Isenberg, 2017), where observations are at the student–teacher link level, and are weighted differentially by teacher dosage. Please see Appendix for more details.
15. Fixed effects for the ETO main effect cannot be included in these models because they would be subsumed by the school fixed effects because they do not vary over time.
16. Prior studies have used similar specifications to directly control for treatment intensity (e.g., Draca, Machin, & Van Reenen, 2011; Duflo, 2001).
17. When investigating whether TFA corps members have differential returns to experience by interacting TFA and years of experience, we generally do not find a differential returns to experience.
18. 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.
19. Yearly grade–subject specific scale score means were retrieved from http://www.fldoe.org/accountability/assessments/k-12-student-assessment/results/.
20. The patterns described in this section continue to hold when analyzing elementary and middle school separately.
21. It is possible that the coefficients reported in Table 5, representing the mean association between density and test scores, mask variation across schools. In an alternative specification, we experimented with interacting density with school fixed effects to obtain school-by-school coefficients. We do not find strong evidence of large across-school variation in effect sizes.
22. Further discussion of the perceived quantity-quality trade-off among district administrators is presented in Hansen et al. (2015).
23. 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.
24. 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.
25. Clark, Isenberg, Liu, and Makowski Zukiewicz (2015) find that TFA teachers outperform non-TFA teachers in a subsample of prekindergarten through Grade 2 teachers, but not in other subsamples.

References

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

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. doi:10.1162/edfp.2006.1.2.176

Bradley, S., Green, C., & Leeves, G. (2007). Worker absence and shirking: Evidence from matched teacher-school data. *Labour Economics*, *14*(3), 319-334. doi:10.1016/j.labeco.2006.05.002

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago: The University of Chicago Press.

Bullough, R. V., Young, J., Erickson, L., Birrell, J. R., Clark, D. C., Egan, M. W., . . . Smith, G. (2002). Rethinking field experience: Partnership teaching versus single-placement teaching. *Journal of Teacher Education*, *53*, 68-80. doi:10.1177/0022487102053001007

Catapano, S. (2006). Teaching in urban schools: Mentoring pre-service teachers to apply advocacy strategies. *Mentoring & Tutoring: Partnership in Learning*, *14*(1), 81-96. doi:10.1080/13611260500432756

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 Zukiewicz, M. (2015). *Impacts of the Teach For America investing in innovation scale-up*. Washington, DC: Mathematica Policy Research.

Collins, J. C. (2001). *Good to great: Why some companies make the leap . . . And others don't*. New York, NY: HarperBusiness.

Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus* (Working Paper No. 17990). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w17990

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

Donaldson, M. L. (2012). The promise of older novices: Teach for America Teachers' age of entry and subsequent retention in teaching and schools. *Teachers College Record*, *114*(10).

Donaldson, M. L., & Johnson, S. M. (2010). The price of misassignment: The role of teaching assignments in Teach For America teachers' exit from low-income schools and the teaching profession. *Educational Evaluation and Policy Analysis*, *32*(2), 25. doi:10.3102/0162373710367680

Donaldson, M. L., & Johnson, S. M. (2011). Teach For America teachers: How long do they teach? Why do they leave? *Phi Delta Kappan*, *93*(2), 47-51. doi:10.1177/003172171109300211

Draca, M., Machin, S., & Van Reenen, J. (2011). Minimum wages and firm profitability. *American Economic Journal: Applied Economics*, *3*(1), 129-151. doi:10.1257/app.3.1.129

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, *91*(4), 795-813. doi:10.1257/aer.91.4.795

Eraut, M. (2004). Informal learning in the workplace. *Studies in Continuing Education*, *26*(2), 247-273. doi:10.1080/158037042000225245

Feeney, N. (2015, August 11). Teach for America welcomes diverse corps after applications drop. *Time Magazine*.

Frank, K. A., Zhao, Y., Penuel, W. R., Ellefson, N., & Porter, S. (2011). Focus, fiddle, and friends experiences that transform knowledge for the implementation of innovations. *Sociology of Education*, *84*, 137-156. doi:10.1177/0038040711401812

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*, 75-96. doi:10.1002/pam.20157

Glazerman, S., Protik, A., Teh, B., Bruch, J., & Max, J. (2013). *Transfer incentives for high-performing teachers: Final results from a multisite randomized experiment* (NCEE 2014-4003). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Hansen, Michael. (2013). Right-Sizing the Classroom: Making the Most of Great Teachers. Thomas B. Fordhom Institute. http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/20131119-Right-Sizing-the-Classroom-Making-the-Most-of-Great-Teachers-FINAL.pdf.

Hansen, M., Backes, B., & Brady, V. (2016). Teacher attrition and mobility during the Teach For America clustering strategy in Miami-Dade County Public Schools. *Educational Evaluation and Policy Analysis*, *38*(3), 495-516. doi:10.3102%2F0162373716638441

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, CALDER Center, Washington, DC.

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. (2017). Methods for accounting for co-teaching in value-added models. *Statistics and Public Policy*, *4*(1), 1-11. doi:10.1080/2330443X.2016.1265473

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. doi:10.1257/app.1.4.85

Johnson, S. M. (2015). Will VAMS reinforce the walls of the egg-crate school? *Educational Researcher*, *44*(2), 117-126.

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. doi:10.1016/j.econedurev.2007.05.005

Kearns, P. (2010). *HR Strategy: Creating business strategy with human capital* (2nd ed.). Oxford, UK: Elsevier.

Koedel, C. (2009). An empirical analysis of teacher spillover effects in secondary school. *Economics of Education Review*, *28*, 682-692. doi:10.1016/j.econedurev.2009.02.003

Kraft, M. A., & Papay, J. P. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*, *36*, 476-500. doi:10.3102/0162373713519496

Lane, S., Lacefield-Parachini, N., & Isken, J. (2003). Developing novice teachers as change agents: Student teacher placements against the grain. *Teacher Education Quarterly*, *30*(2), 55-68.

McDonald, M., & Zeichner, K. (2009). Social justice teacher education. In W. Ayers, T. Quinn, & K. Stovall (Eds.), *Handbook of social justice in education*. Mahwah, NJ: Lawrence Erlbaum.

Papay, J. P., Taylor, E. S., Tyler, J. H., & Laski, M. (2016). *Learning job skills from colleagues at work: Evidence from a field experiment using teacher performance data* (Working Paper No. 21986). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w21986

Penuel, W., Riel, M., Krause, A., & Frank, K. (2009). Analyzing teachers' professional interactions in a school as social capital: A social network approach. *The Teachers College Record*, *111*(1), 124-163.

Rhim, L. M., Kowal, J. M., Hassel, B. C., & Hassel, E. A. (2007). *School turnarounds: A review of the cross-sector evidence on dramatic organizational improvement*. Chapel Hill, NC: Public Impact.

Spillane, J. P., Kim, C. M., & Frank, K. A. (2012). Instructional advice and information providing and receiving behavior in elementary schools exploring tie formation as a building block in social capital development. *American Educational Research Journal*, *49*(6), 1112-1145. doi:10.3102/0002831212459339

Sun, M., Loeb, S., & Grissom, J. A. (2017). Building teacher teams: Evidence of positive spillovers from more effective colleagues.

Educational Evaluation and Policy Analysis, *39*(1), 104-125. doi:10.3102/0162373716665698

Trujillo, T., & Scott, J. (2014). Superheroes and transformers: Rethinking Teach For America's leadership models. *Phi Delta Kappan*, *95*(8), 57-61. doi:10.1177/003172171409500813

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. doi:10.1002/pam.20585

Author Biographies

Ben Backes is a senior researcher at the American Institutes for Research. His recent work has focused on evaluating the training and performance of teachers using state longitudinal databases.

Michael Hansen is a senior fellow at the Brookings Institution and the director of the Brown Center on Education Policy. His research specializes in teacher workforce issues.

Zeyu Xu is a managing researcher at the American Institutes for Research. His research currently focuses on teacher labor market issues and student college readiness.

Victoria Brady is an analyst at TNTP. Her research interests include teacher pipelines and development, education equity, and the sociology of education.