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

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

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

AIR

# How Do Teachers from Alternative Pathways Contribute to the Teaching Workforce in Urban Areas? Evidence from Kansas City

Yang An
Cory Koedel

WORKING PAPER No .243-0920 • September 2020

# How Do Teachers from Alternative Pathways Contribute to the Teaching Workforce in Urban Areas? Evidence from Kansas City

Yang An
*University of Missouri*

Cory Koedel
*University of Missouri/CALDER*

# Contents

## Acknowledgments

We thank the Missouri Department of Elementary and Secondary Education, Teach for America, and Kansas City Teacher Residency for access to data; and Mark Ehlert for research support and feedback on earlier drafts. We gratefully acknowledge financial support from the Ewing Marion Kauffman Foundation and CALDER, which is funded by a consortium of foundations (for more information about CALDER funders, see www.caldercenter.org/about-calder). All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the funders, data providers, or institutions to which the author(s) are affiliated. All errors are our own.



CALDER • American Institutes for Research
1000 Thomas Jefferson Street NW, Washington, DC 20007
202-403-5796 • www.caldercenter.org

***How Do Teachers from Alternative Pathways Contribute to the Teaching Workforce in Urban Areas? Evidence from Kansas City***

Yang An & Cory Koedel
CALDER Working Paper No. 243-0920
September 2020

## Abstract

We examine how teachers from two alternative preparation programs—Teach for America (TFA) and Kansas City Teacher Residency (KCTR)—contribute to the teacher labor market in and around Kansas City, Missouri. We show that TFA and KCTR teachers are more likely than other teachers to work in charter schools, and more broadly, in schools with high concentrations of low-income, low-performing, and underrepresented minority (Black and Hispanic) students. TFA and KCTR teachers are themselves more racial/ethnically diverse than the larger local-area teaching workforce, but only KCTR teachers are more diverse than teachers in the same districts in which they work. In math in grades 4-8 we find sizeable, positive impacts of TFA and KCTR teachers on test-score growth relative to non-program teachers. We also estimate positive impacts on test-score growth in English Language Arts (ELA) for teachers from both programs, but our ELA estimates are smaller in magnitude.

## 1. Introduction

It is well-documented that urban school districts have difficulty recruiting and retaining high-quality teachers (Boyd et al., 2005; Boyd et al., 2006; Lankford, Loeb, and Wyckoff, 2002; Papay et al., 2017). Moreover, recent evidence suggests that accountability policies and improved measures of teaching effectiveness, which have increased the demand for and ability to identify effective teachers, respectively, have exacerbated staffing challenges for high-need schools (Bates, 2020; Cullen, Koedel, and Parsons, forthcoming). Although policy efforts in some states designed to combat these challenges have had some success, teacher recruitment in high-need, urban areas is an ongoing challenge (Glazerman et al., 2013; Springer, Swain, and Rodriguez, 2016; Swain, Rodriguez, and Springer, 2019).

Alternative teacher preparation programs (ATPPs) can be a source of labor supply in localized labor markets that face supply-side challenges. Indeed, many ATPPs explicitly build this idea into their mission statements. A well-known example is the national Teach for America (TFA) program, which we study here. In addition, regionally-based programs with similar goals include New York City Teaching Fellows, the Mississippi Teaching Corps, and Kansas City Teacher Residency (which we also study), among others. Compared to traditional university-based teacher preparation programs, which remain the predominant pipeline into the teaching profession nationally, ATPPs typically provide an accelerated pathway into the classroom. A rationale is that rigid licensing requirements create barriers to entry that keep some qualified teachers out of the classroom (Sass, 2015). By reducing these barriers, ATPPs can increase the appeal and accessibility of the profession for a broader population of potential teachers.

ATPPs also offer pathways to teaching permanency (i.e., paths toward full licensure that would be required for a full career in teaching), although the structure of the pathways differs across programs. There are mixed views about whether ATPPs induce churn in the teaching profession, most notably with respect to TFA, but empirically the evidence suggests the high turnover rate of TFA teachers is not meaningfully different from the rate of other young teachers working in the same challenging environments (Donaldson and Johnson, 2011; Papay et al., 2017). Teacher residency programs, which are an increasingly common form of ATPP, typically include explicit supports to help promote teacher retention and some of these programs produce teachers with much higher retention rates than traditionally-trained teachers (e.g., see Papay et al., 2012).

In this paper, we examine how two ATPPs—TFA and Kansas City Teacher Residency (KCTR)—contribute to the local teacher labor market in and around Kansas City, Missouri. The city school district, Kansas City Public Schools (KCPS), is a high-poverty urban district with low achievement. Surrounding districts are more advantaged than KCPS, albeit marginally in some cases. There is also a large and vibrant charter sector in Kansas City, which is an interesting dimension along which to consider the role of ATPPs in serving the market.

We begin with a descriptive analysis of the TFA and KCTR teacher placements. We show that teachers from both programs are placed disproportionately in charter schools, and more broadly, in schools with larger shares of low-income, low-performing, and underrepresented minority (Black and Hispanic) students. We also examine the diversity of the teachers themselves, motivated by the large minority enrollment share in Kansas City area schools and a rapidly evolving body of research pointing to the importance of demographic representation in the teaching workforce (e.g., Dee, 2005; Egalite and Kisida, 2017; Egalite, Kisida, and Winters, 2015; Holt and Gershenson, 2019; Lindsay and Hart, 2017; Papageorge, Gershenson, and Kang, 2020) . Relative to the larger local area, we find that both TFA and KCTR teachers are more racial-ethnically diverse than other teachers. However, only KCTR teachers are more racial-ethnically diverse than other teachers working in the same districts. TFA, KCTR, and the larger teaching workforce in the Kansas City area are all female-dominated—that said, TFA and KCTR are modestly diversity improving along the dimension of gender.

Next we examine the efficacy of TFA and KCTR teachers as estimated by value added to student achievement in math and English Language Arts (ELA) in grades 4-8. First, for TFA, we estimate that TFA teachers raise student test scores by 0.11 and 0.03 student standard deviations in math and ELA, respectively, compared to non-program teachers on average. These estimates contribute to a large literature on the efficacy of TFA teachers, but to the best of our knowledge are the first estimates from Kansas City. Our findings are consistent with previous evidence that TFA teachers are much more effective than other teachers in similar circumstances in terms of raising math achievement; and either similar to, or marginally more effective than, other teachers in terms of raising ELA achievement.[1]

We are not aware of any previous efficacy evidence for KCTR teachers, for whom our efficacy findings are similar to what we find for TFA. Specifically, we find that KCTR teachers increase student achievement by 0.15 and 0.05 student standard deviations in math and ELA, respectively, compared to non-program teachers on average. Despite strong interest in the teacher residency model among teacher educators (Guha, Hyler, and Darling-Hammond, 2016), our results

[1] This description of the empirical literature on TFA value-added is broadly accurate, although several studies that have been conducted in New York City find null TFA results. For example, Decker, Mayer, and Glazerman (2004) use a within-school randomized research design to study the effects on student achievement in math and ELA of TFA teachers and estimate that TFA teachers raise student achievement by about 0.15 student standard deviations in math relative to control teachers in their same schools. Backes et al. (2019) use value-added models and data from Miami-Dade County and find that TFA teachers outperform other teachers by about 0.10 student standard deviations in math. Xu, Hannaway, and Taylor (2011) study TFA effects on achievement in high school and find that TFA teachers increase math test scores by about 0.13 student standard deviations. Two studies using data from New York City find smaller-to-null TFA effects in math (Kane, Rockoff, and Staiger, 2008; Boyd et al., 2006). In terms of the effects in ELA, TFA value-added is smaller (Backes et al., 2019; Decker, Mayer, and Glazerman, 2004; Kane, Rockoff, and Staiger, 2008), although in high school, Xu, Hannaway, and Taylor (2011) find that TFA teachers have similar effects on math and English tests.

for KCTR contribute to a very thin literature on the efficacy of teachers from residency programs in terms of their ability to improve student achievement. We are aware of just two previous points of empirical evidence. First, Papay et al. (2012) evaluate the Boston Teacher Residency and find negative impacts on student achievement in mathematics, although they find evidence of a positive performance trajectory among these teachers. The other efficacy evidence is from the Memphis Teacher Residency, which is evaluated as part of Tennessee's Report Card on the Effectiveness of Teacher Training Programs (Tennessee Higher Education Commission, 2014). The report presents mixed results for the Memphis Teacher Residency program, although overall the evidence is more positive than negative.

Taken on the whole, our analysis provides an area-level overview of how the TFA and KCTR programs contribute to the teacher labor market in Kansas City, Missouri. We show that these programs are being used to fill teaching needs in generally disadvantaged districts and schools, including charter schools. And at least as measured by achievement impacts, teachers from these programs are more effective than their non-program peers. In a final, supplementary analysis we examine teacher retention among teachers who enter the labor market via these programs compared to non-program teachers. Consistent with the findings from Papay et al. (2012) on the Boston Teacher Residency, we find that early-career retention for KCTR teachers is far above that of other teachers in the same districts. TFA teachers have higher retention after 1 and 2 years of service, but by year 5 are less likely to remain as teachers in the Kansas City area than other local-area teachers.

## 2. Brief Program Descriptions

### *2.1 TFA*

Teach For America (TFA) recruits high-performing college students who commit to teach for two years in a low-income community where TFA has partnered with local school districts. Pre-placement TFA summer training varies by region but typically includes a 5-7 week accelerated training program, which includes teaching practice and coaching, and a 1-2 week regional induction and orientation program. TFA partners with local certification programs to help corps members pursue full teacher certification during their 2-year commitment period. Donaldson and Johnson (2011) find that the majority of TFA teachers continue to teach beyond the 2-year commitment, although the TFA exit rate increases significantly from the second to third year.

### *2.2 KCTR*

KCTR is an urban teacher residency program operating in the Kansas City area. Residents are college graduates who train with a mentor, receive coaching, and enroll in a Master's program through the University of Missouri-Kansas City. KCTR participants earn credit toward their master's degrees and teach four days a week for a full academic year in their mentor's classroom during the program. At the end of the residency year, residents become certified teachers in Missouri and agree to teach in a high-need school in Kansas City for three additional years. During the first post-residency year, program participants complete their Master's degrees, and they continue to receive instructional coaching throughout the three-year post-residency commitment.

## 3. Data

We received comprehensive lists of TFA and KCTR participants placed in Missouri schools from the programs themselves. The data include the year and school of each participant's initial placement after the training. Our TFA data cover seven cohorts who received training between fall-2011 and fall-2017 (inclusive). KCTR is a newer program and the first post-residency cohort was not placed until fall-2017; from KCTR we received program placements for the three cohorts that began their teaching placements in fall-2017, fall-2018, and fall-2019.[2] Hereafter, we refer to each school year by the spring year; e.g., 2017-18 as 2018.

We matched the listed participants to their employment records in administrative data provided by the Missouri Department of Elementary and Secondary Education (DESE). The DESE data provide additional information about the participants themselves, their placements, and their students. We were able to match all of the teachers on the program lists in the DESE data.

Table 1 shows the counts of program participants by the year of the first post-program placement, again noting that school years are denoted by the spring year. For TFA, we use data for teachers who entered the program between 2012 and 2018 for our evaluation. A small number of TFA teachers entered the workforce with a lag, which is why Table 1 shows non-zero TFA placements in 2019 and 2020. KCTR's initial cohort went through residency during the 2017 school year and our analysis is based on program participants whose first post-residency years were in 2018, 2019 and 2020.

Table 1 also shows the numbers of program participants whose first placements were in teaching positions, who are the focus of our analysis. As expected, the vast majority of program participants were placed in teaching positions. Exceptions include a small number of individuals whose initial placements were not in standard teaching roles. For our descriptive analysis we analyze all teaching placements. For our teacher efficacy analysis based on value-added to student achievement, we use teachers of math and ELA in grades 4-8, for which sample details are

[2] To be more precise, we do not treat the during-residence year as a teaching placement. The first KCTR cohort finished the residency year in spring-2017 and was placed in teaching positions in fall-2017.

provided below. The value-added sample includes "self-contained" elementary teachers and subject-specific teachers in higher grades.

## 4. Methodology

### *4.1 Descriptive Labor Market Analysis*

We begin by describing the composition of teachers and their initial teaching placements for each program compared to other public school teachers in the Kansas City area. As a first step, to define the "local labor market area" or "Kansas City area," we retrieved the address of the central office for each local education agency (LEA) operating over the span of our data from 2012-2020 (including six LEA's that were open in at least one of these years, but closed by 2020). Note that LEA's include both traditional school districts and charter school operators, where the LEA is defined at the level of the operator for charter networks with more than one school in the area. For ease of presentation, we use the terms "LEA" and "district" interchangeably in the text.

We define the local labor market area as including all districts with a Kansas City, Missouri address. There are 30 such districts, including charter authorizers. We also include two additional districts with addresses in nearby Independence and Raytown (which are each about 8 miles from central Kansas City). In total, we define the area to include 32 districts, which combine to represent the region of effect for the programs we evaluate.[3]

Table 2 lists the 32 school districts, ordered from highest to lowest by the percentage of local-area non-program teachers employed, shown in the last column of the table. Noting that the vast majority of local-area teachers are non-program teachers, the ordering is essentially by district size. For each focal program, we report the percent of teachers in our sample from that program

[3] We made one exception in our geographic definition of the Kansas City area, which is to exclude Park Hill school district. While Park Hill has a Kansas City address, it is about 13 miles away from central Kansas City and is a highly advantaged school district. Park Hill did not receive any TFA or KCTR teachers during the period we study.

placed in each district. For ease of presentation, the data are aggregated for programs over relevant years in the 2012-2020 range (per Table 1).

The primary takeaway from Table 2 for KCTR and TFA is that they disproportionately place teachers in the central city school district, Kansas City Public Schools (KCPS). Over 60 percent of TFA teachers are placed in KCPS, whereas no other LEA has a double-digit share of TFA teachers. KCTR's representation in KCPS is also large—it accounts for about 27.6 percent of KCTR placements—but smaller than for TFA. Other districts with double-digit shares of KCTR teachers include Hickman Mills and the network of Crossroads Charter Schools. North Kansas City is the largest school district in the region (based on total enrollment and workforce size), but employs relatively few program teachers, all from KCTR. The North Kansas City student population is much wealthier than the neighboring KCPS population and has a lower share of underrepresented minority (URM; i.e. Black and Hispanic) enrollment.

We compare the composition of teachers and their placements from each program to teachers in the larger Kansas City area in terms of (a) the sector (charter or not), level (elementary, middle/junior high, or high school), and subject of the placement, (b) the characteristics of students in the school, and (c) teachers' own race/ethnicities and genders. Each program is compared to the local area using two different benchmarks. First, we use a simple teacher-weighted average from all 32 districts listed in Table 2 over the years 2012-2020 as a common benchmark for both programs. Second, we construct program-specific benchmarks calculated as district-by-year weighted averages that are unique to each program, where the district-by-year weights are the program-specific teacher shares of initial placements.

Formally, the district-by-year-weighted benchmark value of characteristic X for program $j$, which sends teachers to Kansas City area districts $k$ in years $t$, can be written as:

$$\bar{X}_j = \sum_{n=kt}^{N_{kt}} w_{jkt} X_{kt} \tag{1}$$

In the equation, the weighting variable $w_{jkt}$ is the fraction of all teachers produced by program *j* who are placed in district *k* in year *t*, and $X_{kt}$ is the value of characteristic X for district *k* in year *t*. $N_{kt}$ is the total number of district-by-year cells in which a teacher from any of the two focal programs is placed. In district-years when no teacher from program *j* is placed, $w_{jkt} = 0$. For each program *j*, $\sum_{kt}^{N_{kt}} w_{jkt} = 1$.

The first benchmark, to the simple average over all teachers in the Kansas City area, compares teachers from each program to the region as a whole. The second benchmark, using the district-by-year weights, compares teachers from each program to other teachers in the same districts and years in which teachers from that program are placed. Both are useful for understanding the ways in which the programs influence the regional labor market.

*4.2 Efficacy Analysis*

We estimate the effects of teachers from each program on student achievement in grades 4-8 in math and ELA, on average, compared to non-program teachers during the period 2012-2019 using the following value-added model, structured based on Koedel, Mihaly and Rockoff (2015):

$$Y_{igmpqt} = \beta_0 + \mathbf{Y_{imt\text{-}1}}\boldsymbol{\beta_1} + \mathbf{X_{it}}\boldsymbol{\beta_2} + \overline{\mathbf{Y}}_{\mathbf{mpt\text{-}1}}\boldsymbol{\beta}_3 + \overline{\mathbf{X}}_{\mathbf{pt}}\boldsymbol{\beta_4} + \mathbf{T_{it}}\boldsymbol{\beta_5} + \mathbf{P_{iqt}}\boldsymbol{\beta_6} + \gamma_g + \delta_t + \varepsilon_{igmpqt} \tag{2}$$

In equation (2), $Y_{igmpqt}$ is a standardized test score (standardized by grade-subject-year) for student *i* in grade *g* and subject *m,* who attended school *p* and was taught by teacher *q* in year *t*.[4] $\mathbf{Y_{imt\text{-}1}}$ is

[4] Some students take the algebra-I end-of-course test in the eighth grade instead of the standard grade-level test. We include these students in the analysis and their scores on the algebra-I test are separately standardized.

a 4-element vector of lagged test-score information. The first element is the same-subject lagged score, which we require of all students for inclusion in each subject-specific model (i.e., math or ELA). The second element is the lagged off-subject score—in our models of math achievement we include the lagged ELA score, and in the ELA model we include the lagged math score. To facilitate the inclusion of students who are missing just the off-subject lagged score (but still have the required same-subject score), we impute the missing score to the mean and include an indicator variable that we set equal to one if the score is missing. Finally, we add an interaction between the missing indicator variable and the lagged same-subject score, which improves estimation efficiency by allowing the model to rely more heavily on same-subject lagged performance to predict current performance for students who are missing the off-subject lagged score. The vector $\bar{\mathbf{Y}}_{\mathbf{mpt\text{-}1}}$ includes school-average values of the lagged test-score variables (lagged math achievement, lagged ELA achievement, and the fraction missing the off-subject test).

The vector $\mathbf{X}_{\mathbf{it}}$ contains student characteristics. We include indicators for racial/ethnic and gender designations, free and reduced-price lunch (FRL) status, individualized education program (IEP) status, English language learner (ELL) status, and mobility status (i.e., an indicator for whether the student changed schools mid-year during year *t*).[5] We also include school percentages of these variables in the vector $\bar{\mathbf{X}}_{\mathbf{pt}}$.[6] The vector $\mathbf{T}_{\mathbf{it}}$ controls for teacher experience. In our preferred specification we bin teachers into experience groups as in Clotfelter, Ladd, and Vigdor (2007): (1) 0 years prior experience (omitted category), (2) 1-2 years, (3) 3-5 years, (4) 6-12 years, (5) 13-20 years, (6) 21-27 years, and (7) 28+ years. The inclusion of the experience bins ensures

[5] The racial-ethnic categories we include are American Indian, Asian/Pacific Islander, Black, Hispanic, and multi-race (White is the omitted group).
[6] For parsimony we condense the racial-ethnic school percentage variable to capture just the percentage of non-White, non-Asian/Pacific Islander students; this has no substantive effect on our results.

teacher comparisons are restricted to occur within these experience bands. We also estimate a version of the model where we omit teacher experience entirely.[7] We elaborate below on the insights afforded by the comparison of models with and without experience controls. $\gamma_g$ and $\delta_t$ are grade and year fixed effects, respectively, and $\varepsilon_{igmpqt}$ is the error term, which we cluster at the teacher level following Koedel et al. (2015). The vector $\mathbf{P_{iqt}}$ includes the treatment variables of interest: two separate indicator variables for whether student *i*'s teacher in year *t* is from one of the focal programs. The omitted comparison group consists of non-program teachers in the Kansas City area.[8]

We estimate value-added for the subsample of teachers in grades 4-8 in math and ELA. We exclude data from 2020 from the value-added analysis because like other states, Missouri halted 2020 testing due to the Covid pandemic (this has a disproportionate effect on our KCTR sample, per Table 1). In math, our value-added analysis includes 146 TFA teachers and 20 KCTR teachers. In ELA, the analogous teacher sample sizes are 147 and 24, respectively (some of these are overlapping—i.e., self-contained teachers in elementary schools). Due to the clustering structure of the models, the teacher sample sizes are the key determinants of statistical power. Our large TFA sample allows for fairly precise inference regarding program-level value-added. Our

[7] In results omitted for brevity we also confirm that all of our main findings are qualitatively upheld if we control for experience using linear and quadratic terms in place of the bins.

[8] The comparison group includes teachers from a subset of the 32 LEAs listed in Table 2. This is because a few LEAs do not cover grades 4-8 during the period 2012-2019 (e.g., a K-3 charter school). A small data issue also arises for teachers that start in one of the 32 focal districts, but subsequently move to a different district. For our main models, we include all teacher-year observations in any of the 32 focal school districts, and drop all observations outside of these districts (e.g., when a teacher moves out of the area but remains in Missouri). That said, how we handle data outside of the 32 focal districts is inconsequential to our results. For example, we have confirmed that the value-added results are qualitatively insensitive to including teacher-years from outside districts when teachers move. We have also confirmed our results are similar if we pull in more control teachers from the districts that teachers move to from our original sample. The robustness of our results to modifying the sample-inclusion criteria is consistent with the model's ability to control for student and school circumstances to isolate teacher effects on student learning (Koedel, Mihaly, and Rockoff, 2015).

standard errors for the KCTR estimates are larger (about 60-100 percent larger depending on the outcome); they are still informative, but future research on KCTR (and other teacher residency models) would benefit from analyses at greater scale. A general challenge is that the scale of teacher residency programs is often modest, especially when one accounts for the fact that not all teachers are placed in tested grades and subjects. For example, in Papay et al. (2012)—the only other published study we are aware of focused on a teacher residency program that estimates value-added—their sample of residency teachers is similarly modest in size (N≈50).

## 5. Results

### *5.1 Descriptive Analysis*

Figures 1-4 document the compositions of program teachers along the dimensions of school type and level of placements, characteristics of students at placement schools, and the demographics of teachers themselves. The figures are structured so that there is one graph for each program in each figure. For a given characteristic, the blue bars show average values for teachers in the focal program. The orange bars show average values for non-program teachers in the local area—i.e., the simple averages across all non-program teachers in the districts listed in Table 2. The grey bars show district-by-year weighted average values for non-program teachers as calculated by equation (1). Note that the non-program group excludes teachers from both focal programs to facilitate its consistency across comparisons. In the appendix, we provide data tables with all of the information presented in the figures (Appendix Tables A1-A3). In addition, the appendix tables show comparisons restricted to only novice teachers (0-2 years) and provide some additional details that we omit from the figures for ease of exposition.

We illustrate the substantive difference between the orange and grey bars in the figures using TFA as an example. Returning to Table 2, note that KCPS employs 19.14 percent of all non-program local-area teachers, and thus the orange bars in our comparisons involving TFA (implicitly) give a 19.14 percent weight to KCPS when setting the comparison group. However, Table 2 also shows that TFA places a disproportionate fraction of teachers in KCPS—specifically, 60.58 percent of TFA teachers are initially placed in KCPS. The grey bars re-weight KCPS so that it has a 60.58 percent weight in the TFA-specific comparison group. In other words, the orange bars compare TFA teachers to the local-area average on the whole, whereas the grey bars compare TFA teachers to other teachers in the districts (and years) that match TFA's own placement profile.

Beginning with Figure 1, we document teacher placements in terms of the schooling level and whether the placement is in a charter or non-charter school. We use DESE's rules to categorize each school as either an elementary school, middle/junior high school, or high school, as follows: Elementary schools are defined as schools with any combination that includes grades K-8, middle schools are those with any combination that includes grades 4-8 and is at least partly departmentalized, junior high schools have any combination that includes departmentalized grades 7-9, and high schools typically include grades 9-12 but may include grades 7-12.

For TFA, Figure 1 indicates that about 48 percent of teachers in our sample were placed in elementary schools. This value is below the simple average of the local area, which is about 56 percent, and also below the TFA-specific weighted average comparison group in the same districts and years, which is about 55 percent. Thus, from the first set of bars we conclude that TFA teachers are less likely than other teachers in the local area, and other teachers in the same districts and years in which TFA placements occur, to teach in elementary schools. The graph shows that the underrepresentation of TFA teachers in elementary grades is made up in high schools, where TFA teachers are disproportionately likely to be placed. In contrast, for KCTR teachers, the figure shows that they are more likely to be placed in elementary schools but less likely to be placed in high schools relative to the larger local-area labor market.

The final set of bars in each graph in Figure 1 shows the percentages of teachers across all schooling levels placed in charter schools. Both TFA and KCTR teachers are much more likely to teach in charter schools than the average non-program teacher, which highlights the charter sector's disproportionate reliance on these programs for staffing. As indicated by the blue bars, the charter percentages for TFA and KCTR are about 38 and 49.5 percent, respectively. In comparison, as indicated by the orange bars, just 15.5 percent of non-program teachers in the area teach in charter schools. For the charter school comparison in particular, the weighted average comparisons given by the grey bars are not especially informative because almost all charter school operators are coded as their own districts in Missouri. Because the weighted-average comparison group forces weights proportional to each program's own district placements, it is by construction that the percentage of teachers in charter schools for each program virtually matches its program-specific weighted-average value indicated by the grey bar.

Next, in Figure 2 we document average student characteristics at teachers' placement schools. The structure of the figure is the same as Figure 1. We compare program teachers' placement schools using four school-level student characteristics: (1) the underrepresented minority (URM) enrollment percentage, which we calculate as the percentage of Black and Hispanic students (note that given the demographics of the local area, the URM percentage primarily captures the percentage of students who are Black), (2) the free and reduced-price lunch (FRL) eligible enrollment percentage, (3) the percentage of students on an individualized education program (IEP), (4) the percentage of students who are English Language Learners (ELL).[9]

In terms of demographics, Figure 2 shows pronounced differences in the URM percentages between program and non-program teachers in the local area for TFA and KCTR—program teachers are much more likely to work in schools with higher URM student populations than non-program teachers. This can be seen by the gap between the blue and orange bars in each graph corresponding to the URM percentages. Note that for both programs, the URM-percentage gap disappears when the comparison shifts to the program-specific weighted averages, represented by the grey bars. This is informative about the mechanism by which program teachers are disproportionately working with high URM populations. Specifically, it means that the sorting is all occurring at the district level. Put another way, conditional on the district and year of the placement, teachers from the focal programs are working in schools with similar URM percentages as other, non-program teachers. But because the districts in which they are placed are high-URM districts, their exposure to URM students is higher than the local-area average of all teachers.

With respect to student FRL and ELL percentages, a qualitatively similar pattern plays out for each program, with modest variability in the magnitude of exposure gaps between program and non-program teachers. There is no indication of differences in schools' IEP percentages for program and non-program teachers.

Figure 3 provides related evidence using average standardized test scores at teachers' placement schools. We make two notes about these comparisons. First, average test scores in all teachers' schools, even in the larger sample of non-program teachers, are large and negative in

[9] The FRL percentage is measured imperfectly because of the community eligibility provision, or CEP (Koedel and Parsons, 2020). The CEP-induced measurement error in the FRL percentage is likely to understate differences between program and non-program teachers along this dimension, but directionally the comparisons are still informative.

both subjects. This is because we standardize scores using the state distribution. The implication is that on average, students in the Kansas City area (as we have defined it) perform below the state average. Second, the test score results in Figure 3 are descriptive only. They may embody some program effects to the extent that program teachers impact test scores, about which we provide some evidence below. However, noting that non-schooling factors explain the majority of the variance in student test score levels (Parsons, Koedel, and Tan, 2019), and program teachers represent just a small fraction of the local-area workforce, our primary use of school-average test scores here, like in the previous figures, is to provide information about placement context.

Figure 3 shows that both TFA and KCTR teachers are placed in schools with substantially lower test score levels than other local-area teachers, as indicated by the large gaps between the blue and orange bars for these programs in Figure 3. Like the comparisons using the other student characteristics in Figure 2, the gaps shrink when we use the program-specific weighted averages, although they do not completely close as was the case for the previous comparisons. The fact that they mostly close points to district placements, and not school placements within districts, as the primary mechanism that drives the sorting of TFA and KCTR teachers into schools with lower achievement. But the fact that the gaps do not close all the way indicates that there is some additional within-district sorting of TFA and KCTR teachers that leads them to teach at lower-achieving schools compared to other teachers in the same districts and years. In the appendix (Appendix Tables A2.b and A2.c), we show that the small gaps that remain are partly explained by teacher experience. Specifically, noting that TFA and KCTR teachers are themselves inexperienced, if we restrict the weighted comparison group for each program to include only inexperienced non-program teachers, the achievement gap between program and non-program teacher placements declines further.[10]

We additionally note one other finding from the achievement comparisons not shown in Figure 3, but available in the appendix: the sorting patterns of program teachers in tested grades and subjects largely mirrors the sorting patterns of all program teachers discussed thus far. This is important for informing our analysis of value-added. If program teachers in tested grades and subjects are sorted to schools differently from other program teachers, it could have implications for inference from the value-added models and how the results relate to the descriptive comparisons. However, Appendix Tables A2.a-A2.c show that the sorting patterns are similar, which aids in the interpretation of the value-added results.

Finally, Figure 4 documents the racial-ethnic and gender compositions of program teachers themselves relative to the local area using the same structure as the previous figures. We divide teachers into the following racial/ethnic groups: Asian/Pacific Islander, Black, Hispanic, White, and Other. The Other category is suppressed for ease of presentation, but results are reported in the appendix (Appendix Tables A3.a-A3.c). Compared to local area teacher demographics overall, as represented by the orange bars in the graphs, both programs are at least modestly diversifying, with generally larger percentages of Asian/Pacific Islander, Black, and Hispanic teachers, and smaller percentages of White teachers, than the local-area average. KCTR is the most diverse program, particularly with respect to the percentage of Black teachers (37 percent, which is about 2.5 times higher than the local-area average of 13.55 percent). When we compare program teachers to the program-specific, district-and-year weighted average comparison groups (grey bars), KCTR

[10] This is consistent with well-documented evidence that inexperienced teachers, on average, are more likely to work in disadvantaged schools (Clotfelter, Ladd, and Vigdor, 2006; Goldhaber, Quince, and Theobald, 2018).

teachers remain more diverse racial-ethnically than their non-program teaching peers; however, TFA teachers are a less diverse group.

We also examine gender diversity. Like the national teaching workforce, the workforce in the Kansas City area is female-dominated, as is each focal program. That said, the both TFA and KCTR are modestly diversity-improving along the dimension of gender.

*5.2 Efficacy Analysis*

Figure 5 shows the main value-added results for teachers in grades 4-8. We estimate four different models for each subject—all variants of equation (2)—to recover estimates of the average value-added of teachers from each program relative to non-program teachers. A solid bar in the figure indicates the estimate is statistically distinguishable from the average value-added of non-program teachers at the 10 percent level or better and a clear bar indicates the estimate cannot be distinguished from the value-added of non-program teachers. The results underlying the figure are also available in tabular form in Appendix Tables A4 (math) and A5 (ELA).

The four different value-added specifications are labeled as models 1-4 in the figure. Model 4 is the full specification shown in equation (2) and models 1-3 are sparser variants that build up to the full model. First, model 1 is a base specification that only controls for individual lagged achievement $(\mathbf{Y}_{\mathbf{imt-1}})$ and grade and year fixed effects. Model 2 adds the individual student characteristics in the vector $\mathbf{X}_{\mathbf{it}}$. Model 3 further adds the school-level control vectors $\bar{\mathbf{Y}}_{\mathbf{mpt-1}}$ and $\bar{\mathbf{X}}_{\mathbf{pt}}$ to account for schooling context. The last component of the full model is the vector of teacher experience bins, denoted by $\mathbf{T}_{\mathbf{it}}$, which is added to model 4.

Before describing the results, we first note that our value-added estimates reflect the combined effects of (a) any selection into the programs and (b) any incremental improvement in teaching caused by the programs conditional on who enrolls. A program can have high value-added through either or both channels. For example, if a program recruits individuals who are predisposed to be strong teachers (i.e., positive selection) but does nothing via training to improve their performance, it will have high value-added; similarly, a program that recruits average teachers but offers exemplary training will also have high value-added. While our inability to disentangle the "selection" and "training" effect mechanisms is a limitation for some research questions, the combined effect is likely of first-order policy interest for districts looking to hire effective teachers.[11]

We first focus on the results from model 4, which is our preferred specification because it controls for student and school circumstances and compares teachers with similar experience levels. In math, the findings indicate that TFA and KCTR teachers outperform non-program teachers in the Kansas City area. Their value-added estimates are 0.11 and 0.15 student standard deviations higher on average, respectively. To give these estimates some context, the best research on teacher quality indicates that a one-standard-deviation move in the distribution of teacher quality as measured by math value-added—e.g., a move from about the 50$^{th}$ to 85$^{th}$ percentile in the distribution of teacher quality—corresponds to a move in the student test distribution of about 0.10-0.15 standard deviations. Thus, our estimates of 0.11 and 0.15 imply that TFA and KCTR teachers are about 0.70-1.50 teacher standard deviations more effective than comparable non-program teachers in our data, on average. These are very large effects.

Comparing our findings to previous studies on the value-added of TFA teachers in math, our positive estimate for TFA is qualitatively consistent with the literature outside of the New York City (e.g., Boyd et al., 2006; Kane, Rockoff, and Staiger, 2008), and inclusive of the New York

[11] That is, districts will care more about whether effective teachers come out of a particular pipeline than why one pipeline produces stronger teachers than another. Disentangling the mechanisms is of greater interest from the perspective of informing teacher training organizations, for which knowing more about what aspects of training lead to greater improvements in efficacy for the teachers who participate is important.

City estimates, falls somewhere in the middle of the range of estimates in previous research.[12] We are not aware of any comparable prior literature on KCTR specifically. However, a similar evaluation of the efficacy of teachers from the Boston Teacher Residency (BTR) by Papay et al. (2012) finds negative achievement effects in math.

Looking at the estimates for TFA and KCTR across models in Figure 5 is also instructive. In the sparse model—model 1—there are no statistically detectable differences between the program and non-program teachers. However, once we control for student characteristics in model 2, the differences emerge and persist as the specifications become richer. This finding is previewed by the descriptive analysis above, which shows that TFA and KCTR teachers are more likely to be placed in schools with more disadvantaged and lower-achieving students. Model 1 does not account for these placement differences except to the extent that they are captured by students' own lagged test scores. The more robust accounting for teaching context in models 2-4 reveals important performance gaps between TFA and KCTR teachers compared to other Kansas City area teachers.

Another aspect of the cross-model estimates that merits attention is the difference between models 3 and 4. These models differ only by whether we control for teacher experience. On the one hand, the experience-conditional comparisons in model 4 are useful for gauging the efficacy of TFA teachers relative to their similarly-experienced non-program peers. However, it is also desirable to compare TFA teachers to all program teachers without conditioning on experience because part of the TFA treatment, arguably, is increased student exposure to relatively inexperienced teachers. Model 3 does not separately control for experience, so implicitly it compares TFA teachers to all non-program teachers, who are much more experienced on average.

[12] Again, these studies include Decker, Mayer, and Glazerman (2004), Backes et al. (2019), and Xu, Hannaway, and Taylor (2011).

The estimates from both models are informative about the TFA treatment effect. We have less information on the longevity of KCTR teachers and no *ex ante* reason to believe they will have shorter careers, on average, than non-program teachers, so for KCTR teachers the comparisons in model 4 are preferred.

As a practical matter, our estimates change little going from models 3 to 4 for both TFA and KCTR. Upon further investigation, the reason for the modest change in the estimates is that the experience-efficacy gradient among non-program local area teachers is modest, and effectively flat over a large range of the experience distribution after the first year (results suppressed for brevity). While this result is not entirely out of line with what has been found elsewhere in the literature (Clotfelter, Ladd, and Vigdor, 2006; Wiswall, 2013), the gradient in Kansas City is especially flat.[13] Thus, whether we condition on teacher experience in our comparisons involving TFA and KCTR teachers has little bearing on the findings.

Next, we turn to the ELA estimates in the second panel of Figure 5. The results from model 4 suggest small positive effects of TFA and KCTR teachers on ELA growth, compared to similarly-experienced non-program teachers, on the order of about 0.03-0.05 student standard deviations. Our smaller findings in ELA are consistent with the broad empirical regularity documented in research that teacher effects in math are larger than in English Language Arts (e.g., Lefgren and Sims, 2012; Goldhaber, Cowan, and Walsh, 2013). For TFA specifically, these findings also align with findings from related studies (e.g., Backes et al., 2019; Decker, Mayer, and Glazerman, 2004; Kane, Rockoff, and Staiger, 2008).[14] The ELA effects are not statistically

[13] Wiswall (2013) shows that one explanation for the generally flat, or only slightly upward sloping experience-performance gradient, is negative selection into who stays in teaching. That said, for our purposes the distinction of mechanisms is not critical.

[14] Decker, Mayer, and Glazerman (2004) and Kane, Rockoff, and Staiger (2008) find no statistical evidence of a TFA effect in ELA, although their standard errors also cannot rule out modest positive impacts, and Backes et al. (2019) estimate a statistically significant TFA effect on ELA of 0.02 student standard deviations (which is very close to our estimate).

preserved in model 3, where we do not condition on teacher experience, although the absolute impact of going from model 3 to model 4 is similar in ELA and math. The difference is that the estimates from model 4 for ELA are already small, so the modest reduction in the coefficients going to model 3 pushes them below the significance threshold.

**6. Extension: Teacher Retention**

We briefly extend our analysis to assess whether program teachers are more or less likely to remain in the Kansas City area compared to non-program teachers. For KCTR we can perform the retention analysis for the 2018 placement cohort only—the 2019 and 2020 cohorts are too new to credibly study retention. We look to see if the 2018 cohort of KCTR teachers remain in the workforce in 2019 and 2020 (N=31, noting that one 2018 placement was in a non-teaching position). Because we have more cohort data for TFA, we expand the retention analysis to look forward up to five years for TFA teachers whose initial placements were between 2012-2016 (inclusive; N=340).

Retention rates in the Kansas City area for both programs are reported in Figure 6 compared to retention rates for non-program first-year teachers in the same years. We define the Kansas City area more broadly for examining retention than in the previous analysis. Specifically, we treat a teacher as retained in the Kansas City area if she is observed teaching in a public school in the Missouri portion of the formal metropolitan statistical area (MSA) as defined by the U.S. Census. We also report retention rates in the larger Missouri workforce in the appendix (which are slightly higher but similar; see Appendix Tables A6 and A7).

In Figure 6, retention after one year indicates that the teacher was observed working in the year following the initial placement (i.e., in year 2). Retention rates after years 2, 3, and 4 are similarly defined and cumulative (e.g., a value of 50 after year-3 would indicate that 50 percent of the entering teachers were still teaching in the area into the 4th year). As in previous figures, we compare KCTR and TFA teachers to the simple average of teachers in the districts listed in Table 2, and the program-specific weighted averages based on the districts and years in which KCTR and TFA teachers were placed. We restrict the comparison groups to new teachers only for this analysis.

Figure 6 shows that 1- and 2-year retention rates for KCTR are above the local-area average and the district-by-year weighted average based on KCTR placements. The retention gaps over the first two years for KCTR teachers are quite large—after one year, KCTR teachers are more than 20 percentage points more likely to remain in teaching in the area relative to non-program new teachers in the same districts and years. After two years the gap shrinks but remains sizeable, at about 14 percentage points. The higher retention rates of KCTR teachers over the span of time we can evaluate is consistent with the feature of the program that residents agree to teach in a high-need school in the Kansas City area for at least three years.

For TFA, more than 99 percent of TFA teachers in our sample return after the first year, which is consistent with the 2-year program commitment. However, there is a stark drop off going into year 3, with only 57 percent of TFA teachers remaining beyond the second year. Retention after the 4th year—i.e., the percent of TFA teachers who are still teaching in year-5—is just 32.35 percent. These retention rates for TFA teachers in our sample are similar to rates calculated using national TFA data (Donaldson and Johnson, 2011).

The seemingly low retention rate of TFA teachers in the metro area—32.35 percent—is perhaps disappointing, but less so given the larger context that non-program teachers in these same districts also have a low 5-year retention rate in the area (40.84 percent). As noted by Donaldson and Johnson (2011), this points to the issue of high turnover being less about the TFA program per se, and more about the difficult teaching environments faced by TFA teachers. While non-program teachers are retained at a higher rate than TFA teachers in the Kansas City area by the fifth year, the retention gap is not so large as to be a first-order difference between TFA and other teachers.

**7. Conclusion**

We evaluate two alternative teacher preparation programs—Teach for America (TFA) and Kansas City Teacher Residency (KCTR)—to assess how they contribute to the teacher labor market in Kansas City, Missouri. Descriptively, we document program placements in terms of school types and levels, characteristics of students taught, and the racial/ethnic and gender diversity of the teachers themselves. Although there is some heterogeneity across the two programs, common themes are that these programs disproportionately place teachers in charter schools, and more broadly, in schools serving disadvantaged students. Teachers from both programs are also more racial-ethnically diverse than the larger local-area teaching population, although only KCTR teachers are more diverse than other teachers in the same districts in which they place. Notably, KCTR seems to be particularly effective as a pathway for Black teachers to enter the profession.

We also estimate the value-added of program teachers relative to non-program teachers. We find that students in grades 4-8 whose teachers are from TFA and KCTR have much higher achievement growth in math than similar students, in similar schools, who are taught by non-program teachers. We also find evidence of small, positive impacts of teachers from these programs on ELA achievement in grades 4-8.

Finally, we analyze teacher retention for TFA and KCTR teachers. We find that KCTR teachers are much more likely to be retained in Kansas City MSA than comparable non-program teachers over the first 3 years post-training. TFA teachers are more likely be retained after the first year, but their retention rates drop off thereafter. Retention rates after 4 years for TFA teachers are below comparable rates for non-program teachers in the same districts, but not markedly (32 versus 41 percent).

## References


Backes, B., Hansen, M., Xu, Z., and Brady, V. (2019). Examining spillover effects from Teach for America corps members in Miami-Dade County Public Schools. *Journal of Teacher Education* 70(5), 453-471.

Bates, M. (2020). Public and private employer learning: Evidence from the adoption of teacher value added. *Journal of Labor Economics* 38(2), 375-420.

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

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

Clotfelter, C., Ladd, H., and Vigdor, J. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources* 41(4), 778-820.

Cullen, J.B., Koedel, C., and Parsons, E. (forthcoming). The compositional effect of rigorous teacher evaluation on workforce quality. *Education Finance and Policy*.

Decker, P., Mayer, D.P., and Glazerman, S. (2004). The effects of Teach for America on students: Findings from a national evaluation. *Princeton, NJ: Mathematica Policy Research*.

Dee, T.S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review* 95(2), 158-165.

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

Egalite, A.J., and Kisida, B. (2017). The effects of teacher match on students’ academic perceptions and attitudes. *Educational Evaluation and Policy Analysis* 40(1), 59-81.

Egalite, A.J., Kisida, B., and Winters, M.A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review* 45, 44-52.

Glazerman, S., Protik, A., Teh, B. R., Bruch, J., and Max, J. (2013). Transfer incentives for high-performing teachers: Final results from a multisite randomized experiment. NCEE 2014-4004. *National Center for Education Evaluation and Regional Assistance*.

Goldhaber, D., Cowan, J., and Walch, J. (2013). Is a good elementary teacher always good? Assessing teacher performance estimates across subjects. *Economics of Education Review* 36(1), 216-228.

Guha, R., Hyler, M.E., and Darling-Hammond, L. (2017). The teacher residency: An innovative model for preparing teachers. *Learning Policy Institute*.

Goldhaber, D., Quince, V., and Theobald, R. (2018). Has it always been this way? Tracing the evolution of teacher quality gaps in U.S. public schools. *American Educational Research Journal* 55(1), 171-201.

Holt, S.B., and Gershenson, S. (2019). The impact of demographic representation on absences and suspensions. *Policy Studies Journal* 47(4), 1063-1093.

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

Koedel, C., Mihaly, K., and Rockoff, J.E. (2015). Value-added modeling: A review. *Economics of Education Review* 47, 180-195.

Koedel, C., and Parsons, E. (2020). The effect of the community eligibility provision on the ability of free and reduced-price meal data to identify disadvantaged students. *CALDER Working Paper* No. 234-0320.

Koedel, C., Parsons, E., Podgursky, M., and Ehlert, M. (2015). Teacher preparation programs and teacher quality: Are there real differences across programs? *Education Finance and Policy* 10(4), 508-534.

Lankford, H., Loeb, S., and Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis 24*(1), 37-62.

Lefgren, L., and Sims, D.P. (2012). Using subject test scores efficiently to predict teacher value-added. *Educational Evaluation and Policy Analysis* 34(1), 109-121.

Lindsay, C.A., and Hard, C.M.D. (2017). Exposure to same-race teachers and student disciplinary outcomes for black students in North Carolina. *Educational Evaluation and Policy Analysis* 39(3), 485-510.

Papageorge, N., Gershenson, S., and Kang, Kyungmin. (2020). Teacher expectations matter. *Review of Economics and Statistics* 102(2), 234-251.

Papay, J.P., Bacher-Hicks, A., Page, L.C., and Marinell, W.H. (2017). The challenges of teacher retention in urban schools: Evidence of variation from a cross-site analysis. *Educational Researcher* 46(8), 434-448.

Papay, J.P., West, M.R., Fullerton, J.B., and Kane, T.J. (2012). Does an urban teacher residency increase student achievement? Early evidence from Boston. *Educational Evaluation and Policy Analysis 34*(4), 413-434.

Parsons, E., Koedel, C., and Tan, L. (2019). Accounting for student disadvantage in value-added models. *Journal of Educational and Behavioral Statistics* 44(2), 144-179.

Sass, T.R. (2015). Licensure and worker quality: A comparison of alternative routes to teaching. *The Journal of Law and Economics 58*(1), 1-35.

Springer, M.G., Swain, W.A., and Rodriguez, L.A. (2016). Effective teacher retention bonuses: Evidence from Tennessee. *Educational Evaluation and Policy Analysis 38*(2), 199-221.

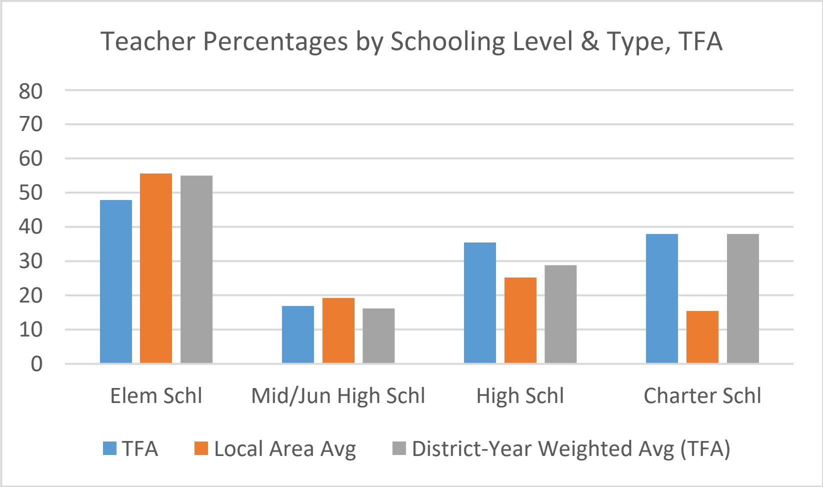
Swain, W.A., Rodriguez, L.A., and Springer, M.G. (2019). Selective retention bonuses for highly effective teachers in high-poverty schools: Evidence from Tennessee. *Economics of Education Review* 68, 148-160.

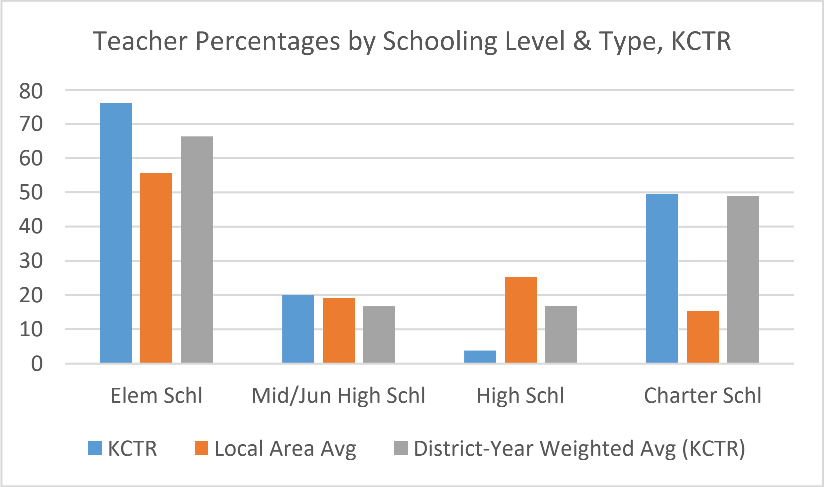
Tennessee Higher Education Commission. (2014). *2014 Report Card on the Effectiveness of Teacher Training Programs*. Tennessee Department of Education.

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics* 100, 61-78.

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

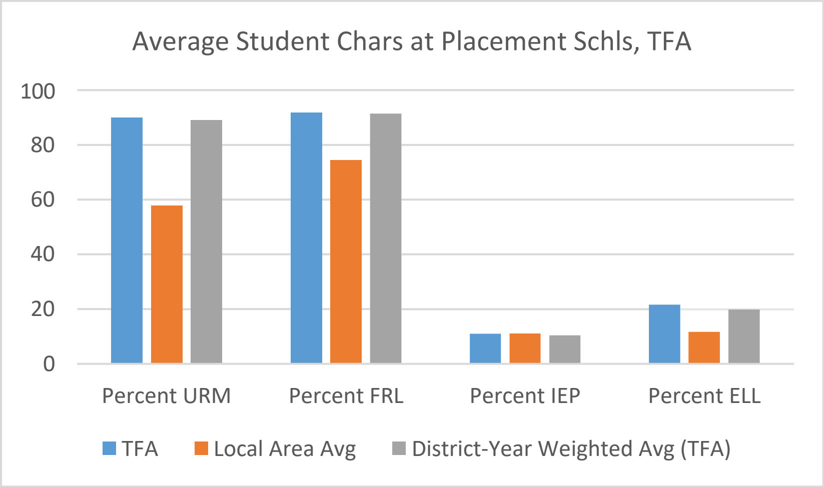
Figure 1. Teacher placements by type school type (i.e., charter or not) and level (elementary, middle/junior high, or high school).

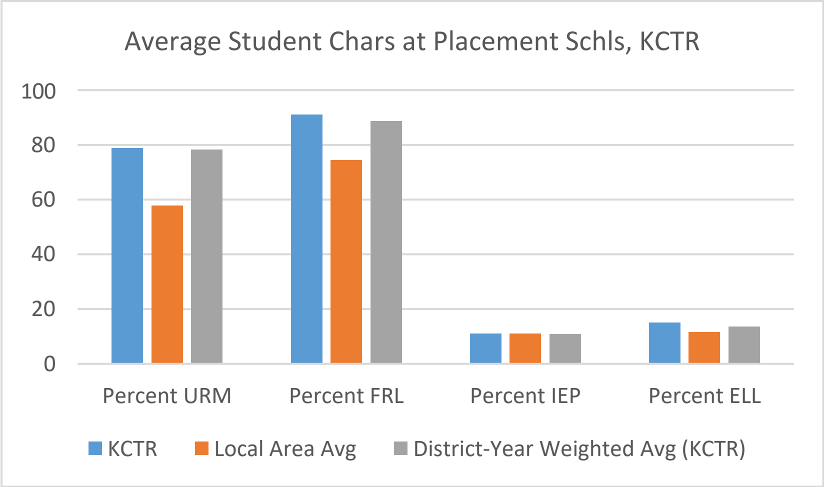



Notes: The local-area averages (orange bars) are for all teachers in the comparison districts shown in Table 2—they are not program specific and thus do not change in the graphs. The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns to compare program teachers to teachers working in the same districts and years of the placements.

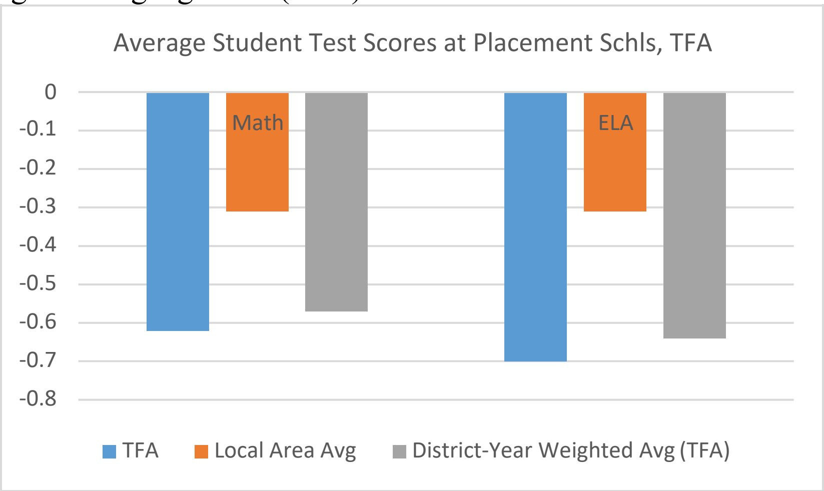
Figure 2. Teacher placements by the characteristics of students attending the initial placement school.

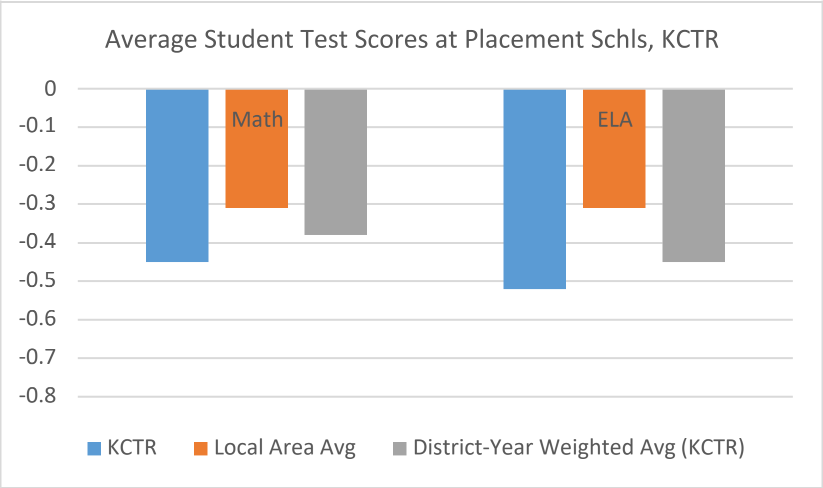



Notes: The local-area averages (orange bars) are for all teachers in the comparison districts shown in Table 2—they are not program specific and thus do not change in the graphs. The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns to compare program teachers to teachers working in the same districts and years of the placements. URM=underrepresented minority (Black and Hispanic); FRL=free or reduced-price lunch eligible; IEP=individualized education program; ELL=English language learner.

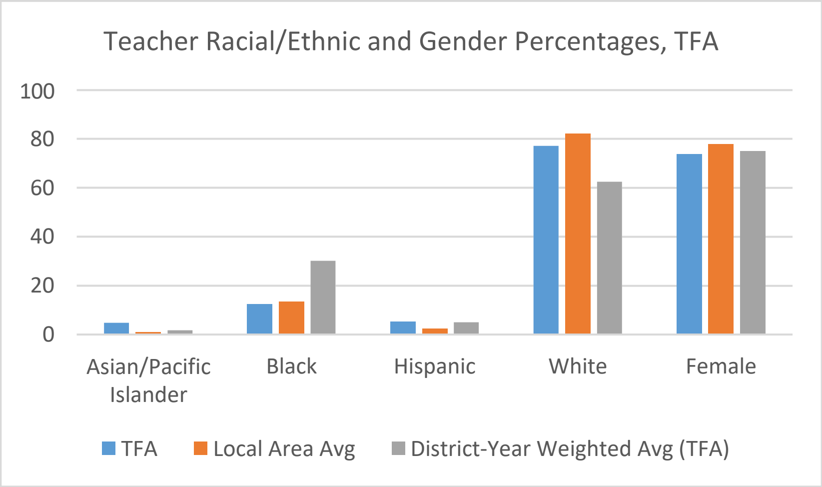
Figure 3. Teacher placements by the standardized achievement level of students attending the initial placement school in math and English Language Arts (ELA).

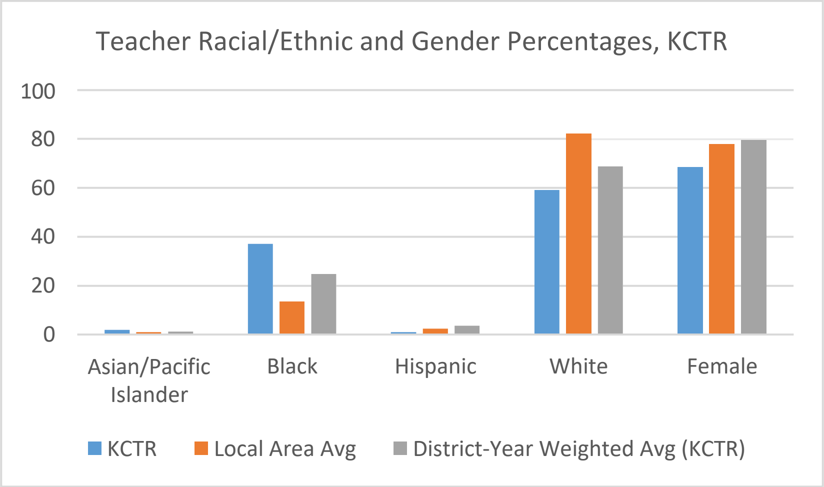



Notes: The local-area averages (orange bars) are for all teachers in the comparison districts shown in Table 2—they are not program specific and thus do not change in the graphs. The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns to compare program teachers to teachers working in the same districts and years of the placements.

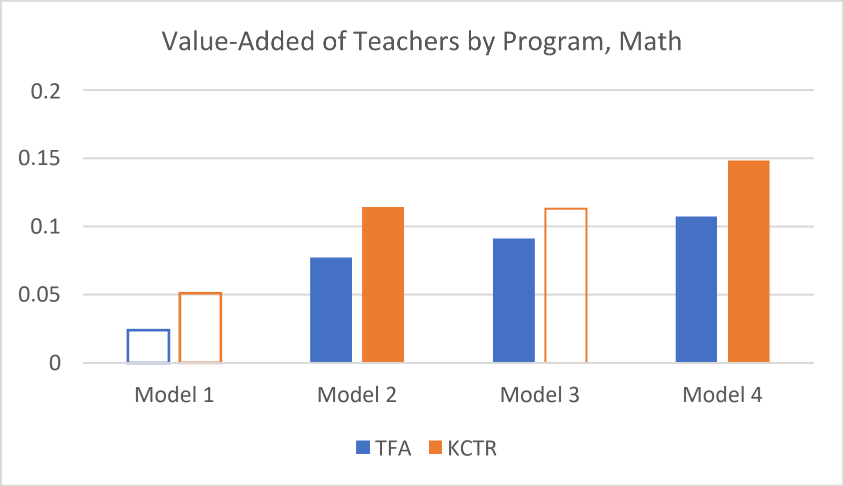
Figure 4. Teachers' racial/ethnic and gender designation percentages.

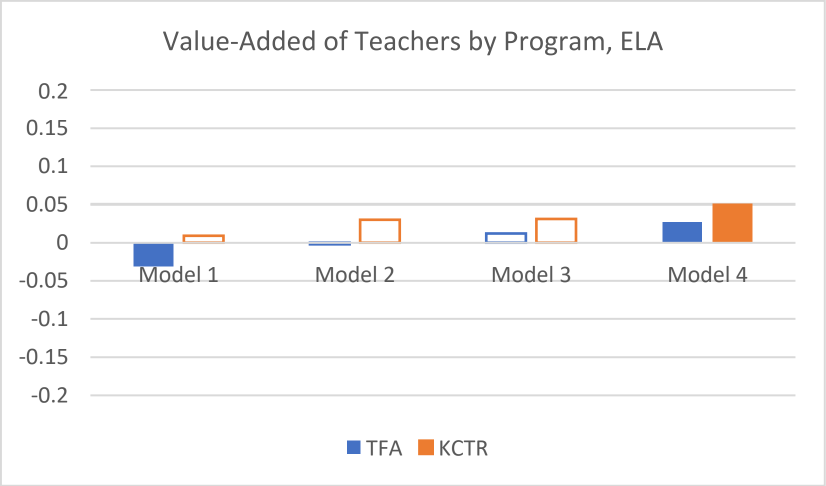



Notes: The local-area averages (orange bars) are for all teachers in the comparison districts shown in Table 2—they are not program specific and thus do not change in the graphs. The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns to compare program teachers to teachers working in the same districts and years of the placements.

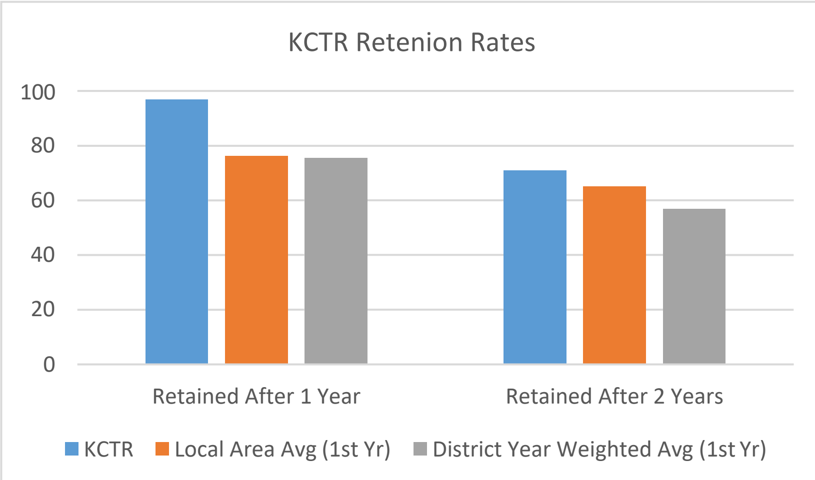
Figure 5. Value-added to achievement in math and English Language Arts (ELA) in grades 4-8 for program teachers compared to non-program teachers using different value-added specifications as described in the text.

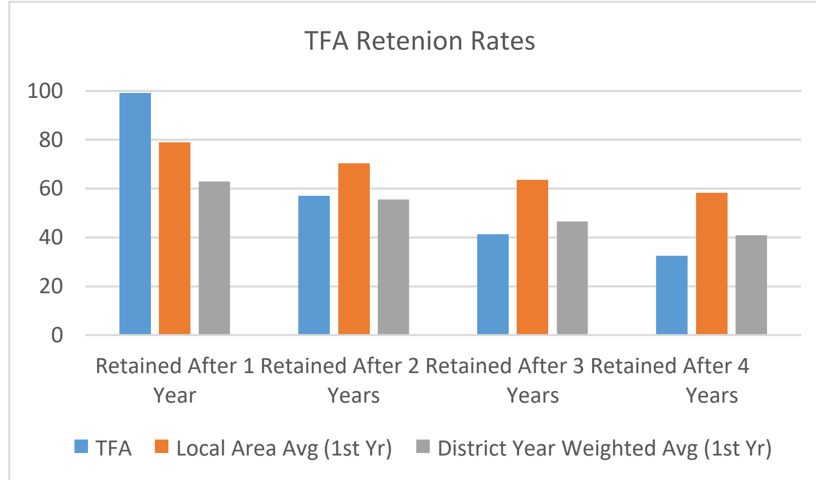



Notes: Moving from model 1 to model 4 increases the comprehensiveness of the value-added model. Model 1 is a base specification that only controls for individual lagged achievement ($\mathbf{Y}_{imt-1}$) and grade and year fixed effects. Model 2 adds the individual student characteristics in the vector $\mathbf{X}_{it}$. Model 3 further adds the school-level control vectors $\bar{\mathbf{Y}}_{\mathbf{pmt-1}}$ and $\bar{\mathbf{X}}_{\mathbf{pt}}$ . The last component of the full model is the vector of teacher experience bins, denoted by $\mathbf{T}_{it}$, added to model 4, which is the full , preferred specification. Solid bars indicate differences that are statistically significant at the 10 percent level or higher; clear bars indicate statistically insignificant differences.

Figure 6. Retention rates for KCTR and TFA teachers relative to novice non-program teachers in the Kansas City Area, as defined by the Missouri portion of the metropolitan statistical area.




Notes: The local-area averages (orange bars) are for all first-year teachers in the full set of comparison districts shown in Table 2, for the same years as program teachers (the definition of this comparison group differs slightly from the definition of the analogous comparison groups above). The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns and thus compare program teachers to other first-year teachers working in the same districts and years of the placements. The retention rates reported in the figure are cumulative (e.g., the "Retained After 3 Years" percentage for TFA reports the number of originally-placed TFA teachers who are still working in the Kansas City MSA in the 4th year after the initial placement). Retention rates for KCTR are reported for the 2018 placement cohort only; retention rates for TFA are reported for the 2012-2016 placement cohorts.

Table 1. Teacher counts by program and the first post-program placement year.

| | TFA[a] | KCTR |
|---|---|---|
| 2012 | 127 | |
| 2013 | 69 | |
| 2014 | 64 | |
| 2015 | 48 | |
| 2016 | 40 | |
| 2017 | 33 | |
| 2018 | 46 | 32 |
| 2019 | 1 | 28 |
| 2020 | 2 | 48 |
| | | |
| Total | 430 | 108 |
| **Total (excluding non-teaching placements)** | **416**[b] | **105**[b] |
| Number unmatched | 0 | 0 |

Notes:
[a] TFA provided placement data for cohorts through 2018. The handful of post-2018 TFA placements are teachers who completed their TFA training in an earlier year but delayed entry into the workforce.
[b] Non-teaching positions include central office positions, individuals listed as working in special centers, and teacher coaches, among other positions.

Table 2. Teacher placement percentages across districts in the Kansas City area, combined across all years.

| | TFA | KCTR | Non-Program Teachers |
|---|---|---|---|
| North Kansas City 74 | 0 | 8.57 | 25.19 |
| Kansas City 33 | 60.58 | 27.62 | 19.14 |
| Independence 30 | 0 | 0 | 16.8 |
| Raytown C-2 | 1.2 | 0 | 10.89 |
| Hickman Mills C-1 | 0.24 | 15.24 | 8.6 |
| Center 58 | 0 | 0 | 4.01 |
| Frontier Schools | 0 | 0 | 2.16 |
| Hogan Preparatory Academy | 6.49 | 0 | 1.39 |
| Academie Lafayette | 0 | 0 | 1.3 |
| University Academy | 4.57 | 3.81 | 1.21 |
| Guadalupe Centers Schools | 6.25 | 0 | 1.14 |
| Kc International Academy | 0.24 | 0.95 | 1.09 |
| Brookside Charter School | 1.92 | 7.62 | 0.81 |
| Lee A. Tolbert Com. Academy | 0.96 | 3.81 | 0.77 |
| Allen Village | 0 | 0 | 0.67 |
| Crossroads Charter Schools | 0 | 10.48 | 0.67 |
| Ewing Marion Kauffman School | 6.25 | 5.71 | 0.57 |
| B. Banneker Academy | 0 | 0 | 0.44 |
| Gordon Parks Elem. | 0 | 0.95 | 0.4 |
| Scuola Vita Nuova | 0 | 1.9 | 0.4 |
| Pathway Academy | 1.2 | 0 | 0.4 |
| Kipp: Endeavor Academy | 2.64 | 7.62 | 0.3 |
| Delasalle Charter School | 1.92 | 0 | 0.29 |
| Genesis School Inc. | 2.88 | 2.86 | 0.24 |
| Derrick Thomas Academy | 1.68 | 0 | 0.22 |
| Hope Leadership Academy | 0 | 0 | 0.2 |
| Renaissance Acad Math and Sci | 0.48 | 0 | 0.18 |
| Academy For Integrated Arts | 0 | 0.95 | 0.18 |
| Citizens Of The World Charter | 0.24 | 0 | 0.17 |
| Hope Academy | 0 | 0 | 0.13 |
| Urban Com. Leadership Academy | 0.24 | 0 | 0.04 |
| Kansas City Girls Prep Academy | 0 | 1.9 | 0.01 |
| | | | |
| Sum | 100 | 100 | 100 |

Notes: Columns sum to 100 percent.

Appendix Tables

Appendix Table A1.a. Teacher placement percentages by grade span and subject for initial post-program placements, combined across all years, compared to the simple region average.

| | TFA | KCTR | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
|---|---|---|---|---|
| *Elementary Total* | 47.84 | 76.19 | 55.53 | 58.67 |
| Tested grades and subjects (4-8) | 12.02 | 28.57 | 10.05 | 11.97 |
| PK-3 | 24.28 | 36.19 | 21.94 | 26.06 |
| Language Specialist | 4.33 | 0.95 | 3.8 | 2.21 |
| Special Education | 4.33 | 0 | 6.15 | 5.97 |
| Other | 2.88 | 10.48 | 13.58 | 12.46 |
| | | | | |
| *Middle School/Junior High School Total* | 16.83 | 20 | 19.19 | 19.18 |
| Tested Grades and Subjects (4-8) | 8.89 | 12.38 | 6.82 | 7.54 |
| Language Specialist | 1.68 | 0.95 | 0.47 | 0.52 |
| Special Education | 1.2 | 0 | 2.15 | 1.5 |
| Science | 2.88 | 3.81 | 2.03 | 2.14 |
| Social Studies | 0.48 | 0.95 | 1.77 | 1.66 |
| Other | 1.68 | 1.9 | 5.94 | 5.82 |
| | | | | |
| *High School Total* | 35.34 | 3.81 | 25.28 | 22.16 |
| Tested Grades and Subjects (7-8) | 7.93 | 0 | 0.56 | 0.55 |
| English Language Arts | 6.25 | 0.95 | 4.05 | 4.15 |
| Math | 4.33 | 1.9 | 2.78 | 2.66 |
| Science | 7.21 | 0.95 | 2.61 | 3.1 |
| Social Studies | 1.92 | 0 | 2.77 | 2.08 |
| Special Education | 4.81 | 0 | 2.93 | 1.97 |
| Other | 2.88 | 0 | 9.58 | 7.64 |
| | | | | |
| Sum (Totals) | 100 | 100 | 100 | 100 |
| | | | | |
| Total Charter Percent | 37.98 | 49.52 | 15.47 | 23.60 |
| | | | | |

Notes: Schooling levels are defined as described in the text. The "other" category at each level contains a number of sparsely populated positions including physical education, health, music, and other specialty subjects and non-traditional assignments. The "total charter percent" row combines charter placements across all schooling levels and subjects. Novice-only teachers in the last column are teachers with 0-2 years of experience.

Appendix Table A1.b. Detailed analog to Table A1.a for TFA, with program-specific weighted-average comparison.

| | | TFA Weighted | |
|---|---|---|---|
| | TFA | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
| *Elementary Total* | 47.84 | 55.03 | 54.62 |
| Tested grades and subjects (4-8) | 12.02 | 9.77 | 9.51 |
| PK-3 | 24.28 | 23.55 | 27.64 |
| Language Specialist | 4.33 | 3.60 | 1.39 |
| Special Education | 4.33 | 6.52 | 5.79 |
| Other | 2.88 | 11.59 | 10.28 |
| | | | |
| *Middle School/Junior High School Total* | 16.83 | 16.14 | 17.77 |
| Tested Grades and Subjects (4-8) | 8.89 | 5.75 | 7.17 |
| Language Specialist | 1.68 | 1.45 | 1.11 |
| Special Education | 1.2 | 1.49 | 0.86 |
| Science | 2.88 | 1.69 | 1.93 |
| Social Studies | 0.48 | 0.96 | 1.74 |
| Other | 1.68 | 4.80 | 4.95 |
| | | | |
| *High School Total* | 35.34 | 28.83 | 27.62 |
| Tested Grades and Subjects (7-8) | 7.93 | 2.34 | 4.18 |
| English Language Arts | 6.25 | 4.12 | 4.09 |
| Math | 4.33 | 2.64 | 2.48 |
| Science | 7.21 | 2.86 | 3.71 |
| Social Studies | 1.92 | 3.60 | 2.26 |
| Special Education | 4.81 | 2.85 | 1.23 |
| Other | 2.88 | 10.42 | 9.67 |
| | | | |
| Sum (Totals) | 100 | 100 | 100 |
| | | | |
| Total Charter Percent | 37.98 | 37.89 | 37.91 |
| | | | |

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by equation (1). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Appendix Table A1.c. Detailed analog to Table A1.a for KCTR, with program-specific weighted-average comparison.

| | | KCTR Weighted | |
|---|---|---|---|
| | KCTR | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
| *Elementary Total* | 76.19 | 66.42 | 67.01 |
| Tested grades and subjects (4-8) | 28.57 | 11.52 | 13.23 |
| PK-3 | 36.19 | 23.70 | 26.41 |
| Language Specialist | 0.95 | 5.15 | 2.83 |
| Special Education | 0 | 7.55 | 7.91 |
| Other | 10.48 | 18.49 | 16.63 |
| | | | |
| *Middle School/Junior High School Total* | 20.00 | 16.74 | 15.62 |
| Tested Grades and Subjects (4-8) | 12.38 | 5.86 | 6.01 |
| Language Specialist | 0.95 | 1.38 | 1.30 |
| Special Education | 0 | 2.28 | 0.92 |
| Science | 3.81 | 1.24 | 1.72 |
| Social Studies | 0.95 | 1.40 | 1.17 |
| Other | 1.90 | 4.58 | 4.50 |
| | | | |
| *High School Total* | 3.81 | 16.84 | 17.37 |
| Tested Grades and Subjects (7-8) | 0 | 0.41 | 0.33 |
| English Language Arts | 0.95 | 2.83 | 3.56 |
| Math | 1.90 | 1.82 | 1.97 |
| Science | 0.95 | 1.89 | 3.62 |
| Social Studies | 0 | 1.79 | 1.41 |
| Special Education | 0 | 1.96 | 1.08 |
| Other | 0 | 6.15 | 5.40 |
| | | | |
| Sum (Totals) | 100 | 100 | 100 |
| | | | |
| Total Charter Percent | 49.52 | 48.85 | 47.45 |
| | | | |

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by equation (1). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Appendix Table A2.a. Average characteristics of students in the schools and years of teachers' first placements, compared to simple average school characteristics (teacher weighted) in the region.

| | TFA | KCTR | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
|---|---|---|---|---|
| Percent Asian/Pacific Islander | 2.40 | 3.31 | 2.91 | 2.71 |
| Percent Black | 64.21 | 62.00 | 39.90 | 45.51 |
| Percent Hispanic | 25.71 | 16.95 | 17.90 | 19.86 |
| Percent White | 6.28 | 12.46 | 33.83 | 27.09 |
| Percent Other | 1.40 | 5.27 | 5.46 | 4.84 |
| | | | | |
| Percent FRL | 91.78 | 90.97 | 74.31 | 80.80 |
| Percent IEP | 10.94 | 10.97 | 11.05 | 11.08 |
| Percent ELL | 21.67 | 15.08 | 11.68 | 14.06 |
| | | | | |
| Average Math achievement (standardized) | -0.62 | -0.45 | -0.31 | -0.41 |
| Average ELA achievement (standardized) | -0.70 | -0.52 | -0.31 | -0.40 |
| | | | | |
| Among teachers in tested grades and subjects only (4-8): | | | | |
| Average Math achievement (standardized) | -0.61 | -0.49 | -0.32 | -0.43 |
| Average ELA achievement (standardized) | -0.69 | -0.54 | -0.31 | -0.41 |

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience.

Appendix Table A2.b. Detailed analog to Table A2.a for TFA, with program-specific weighted-average comparison.

| | | TFA Weighted | |
|---|---|---|---|
| | TFA | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
| Percent Asian/Pacific Islander | 2.40 | 2.46 | 2.11 |
| Percent Black | 64.21 | 65.55 | 69.55 |
| Percent Hispanic | 25.71 | 23.55 | 20.85 |
| Percent White | 6.28 | 6.95 | 5.99 |
| Percent Other | 1.40 | 1.49 | 1.50 |
| | | | |
| Percent FRL | 91.78 | 91.38 | 91.40 |
| Percent IEP | 10.94 | 10.45 | 10.51 |
| Percent ELL | 21.67 | 19.87 | 17.35 |
| | | | |
| Average Math achievement (standardized) | -0.62 | -0.57 | -0.60 |
| Average ELA achievement (standardized) | -0.70 | -0.64 | -0.66 |
| | | | |
| Among teachers in tested grades and subjects only (4-8): | | | |
| Average Math achievement (standardized) | -0.61 | -0.55 | -0.67 |
| Average ELA achievement (standardized) | -0.69 | -0.62 | -0.72 |

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by equation (1). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Appendix Table A2.c. Detailed analog to Table A2.a for KCTR, with program-specific weighted-average comparison.

| | | KCTR Weighted | |
|---|---|---|---|
| | KCTR | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
| Percent Asian/Pacific Islander | 3.31 | 2.71 | 2.66 |
| Percent Black | 62.00 | 61.35 | 62.31 |
| Percent Hispanic | 16.95 | 16.92 | 16.66 |
| Percent White | 12.46 | 14.04 | 13.87 |
| Percent Other | 5.27 | 4.98 | 4.49 |
| | | | |
| Percent FRL | 90.97 | 88.77 | 89.43 |
| Percent IEP | 10.97 | 10.84 | 11.14 |
| Percent ELL | 15.08 | 13.69 | 13.15 |
| | | | |
| Average Math achievement (standardized) | -0.45 | -0.38 | -0.41 |
| Average ELA achievement (standardized) | -0.52 | -0.45 | -0.48 |
| | | | |
| Among teachers in tested grades and subjects only (4-8): | | | |
| Average Math achievement (standardized) | -0.49 | -0.38 | -0.40 |
| Average ELA achievement (standardized) | -0.54 | -0.45 | -0.48 |

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by equation (1). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Appendix Table A3.a. Program teachers' race-ethnicity and gender percentages compared to simple average (teacher weighted) in the region.

| | TFA | KCTR | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
|---|---|---|---|---|
| Percent Asian/Pacific Islander | 4.81 | 1.90 | 1.00 | 1.24 |
| Percent Black | 12.50 | 37.14 | 13.55 | 13.83 |
| Percent Hispanic | 5.29 | 0.95 | 2.40 | 3.18 |
| Percent White | 77.16 | 59.05 | 82.26 | 80.78 |
| Percent Other | 0.24 | 0.95 | 0.79 | 0.97 |
| | | | | |
| Percent female | 73.80 | 68.57 | 77.83 | 77.68 |

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience.

Appendix Table A3.b. Detailed analog to Table A3.a for TFA, with program-specific weighted-average comparison.

| | | TFA Weighted | |
|---|---|---|---|
| | TFA | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
| Percent Asian/Pacific Islander | 4.81 | 1.76 | 2.70 |
| Percent Black | 12.50 | 30.20 | 29.42 |
| Percent Hispanic | 5.29 | 5.00 | 5.79 |
| Percent White | 77.16 | 62.48 | 61.09 |
| Percent Other | 0.24 | 0.56 | 1.00 |
| | | | |
| Percent female | 73.80 | 75.08 | 74.06 |

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by equation (1). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Appendix Table A3.c. Detailed analog to Table A3.a for KCTR, with program-specific weighted-average comparison.

| | | KCTR Weighted | |
|---|---|---|---|
| | KCTR | All Non-Program Teachers | All Non-Program Teachers (Novice Only) |
| Percent Asian/Pacific Islander | 1.90 | 1.25 | 1.77 |
| Percent Black | 37.14 | 24.77 | 25.27 |
| Percent Hispanic | 0.95 | 3.55 | 4.54 |
| Percent White | 59.05 | 68.84 | 67.84 |
| Percent Other | 0.95 | 1.59 | 0.57 |
| | | | |
| Percent female | 68.57 | 79.61 | 75.17 |

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by equation (1). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Appendix Table A4. Value-added to student achievement by program, grades 4-8, math.

| | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| TFA | 0.024<br>(0.036) | 0.077<br>(0.037)** | 0.091<br>(0.038)** | 0.107<br>(0.039)*** |
| KCTR | 0.051<br>(0.069) | 0.114<br>(0.068)* | 0.113<br>(0.069) | 0.148<br>(0.066)** |
| | | | | |
| Lagged test scores, grade & year fixed effects | X | X | X | X |
| Student-level controls | | X | X | X |
| School-level controls | | | X | X |
| Teacher experience controls (bins) | | | | X |
| | | | | |
| R-squared | 0.580 | 0.589 | 0.593 | 0.594 |
| | | | | |
| N (student-year observations) | 185284 | 185284 | 185284 | 185284 |
| N (TFA teachers) | 146 | 146 | 146 | 146 |
| N (KCTR teachers) | 20 | 20 | 20 | 20 |
| N (non-program teachers) | 1953 | 1953 | 1953 | 1953 |

Note: Models control for teacher experience using indicators for the following experience bins as reported in the main text: (1) 0 years prior experience (omitted), (2) 1-2 years, (3) 3-5 years, (4) 6-12 years, (5) 13-20 years, (6) 21-27 years, and (7) 28+ years. Standard errors clustered by teacher are reported in parentheses. The teacher counts reported at the bottom of the table indicate the number of unique teachers (i.e., clusters). *** p<0.01, ** p<0.05, * p<0.10.

Appendix Table A5. Value-added to student achievement by program, grades 4-8, ELA.

| | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| TFA | -0.031<br>(0.014)** | -0.002<br>(0.014) | 0.012<br>(0.013) | 0.027<br>(0.013)** |
| KCTR | 0.009<br>(0.044) | 0.030<br>(0.036) | 0.031<br>(0.026) | 0.051<br>(0.028)* |
| | | | | |
| Lagged test scores, grade & year fixed effects | X | X | X | X |
| Student-level controls | | X | X | X |
| School-level controls | | | X | X |
| Teacher experience controls (bins) | | | | X |
| | | | | |
| R-squared | 0.665 | 0.673 | 0.676 | 0.676 |
| | | | | |
| N (student-year observations) | 186614 | 186614 | 186614 | 186614 |
| N (TFA teachers) | 147 | 147 | 147 | 147 |
| N (KCTR teachers) | 24 | 24 | 24 | 24 |
| N (non-program teachers) | 2178 | 2178 | 2178 | 2178 |

Note: Models control for teacher experience using indicators for the following experience bins as reported in the main text: (1) 0 years prior experience (omitted), (2) 1-2 years, (3) 3-5 years, (4) 6-12 years, (5) 13-20 years, (6) 21-27 years, and (7) 28+ years. Standard errors clustered by teacher are reported in parentheses. The teacher counts reported at the bottom of the table indicate the number of unique teachers (i.e., clusters). *** p<0.01, ** p<0.05, * p<0.10.

Appendix Table A6. KCTR teacher retention rates compared to other teachers in the region.

| | KCTR | All Non-Program Teachers (First-year Teachers Only) | All Non-Program Teachers (First-year Teachers Only, District-year Weighted Average) |
|---|---|---|---|
| Kansas City area 1-year retention rate | 96.77 | 76.24 | 75.48 |
| Kansas City area 2-year retention rate | 70.97 | 64.95 | 56.82 |
| Missouri 1-year retention rate | 96.77 | 78.02 | 76.44 |
| Missouri 2-year retention rate | 74.19 | 67.92 | 58.79 |

Appendix Table A7. TFA teacher retention rates compared to other teachers in the region.

| | TFA | All Non-Program Teachers (First-year Teachers Only) | All Non-Program Teachers (First-year Teachers Only, District-year Weighted Average) |
|---|---|---|---|
| KC area 1-year retention rate | 99.12 | 78.86 | 62.95 |
| KC area 2-year retention rate | 57.06 | 70.35 | 55.51 |
| KC area 3-year retention rate | 41.18 | 63.37 | 46.53 |
| KC area 4-year retention rate | 32.35 | 58.22 | 40.84 |
| MO 1-year retention rate | 99.12 | 80.94 | 65.04 |
| MO 2-year retention rate | 58.82 | 73.27 | 58.08 |
| MO 3-year retention rate | 43.24 | 66.78 | 49.49 |
| MO 4-year retention rate | 33.82 | 61.88 | 43.99 |