ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 22-675

The Effects of High-performing, High-turnover Teachers on Long-run Student Achievement: Evidence from Teach For America

Virginia S. Lovison
Brown University

An increasing share of new teachers enter the profession through alternative certification programs. While these programs increase teacher supply in areas facing critical shortages, they also increase instability in local teacher labor markets via high teacher turnover. A fundamental question is what effect these programs have on student achievement over the long run. To address this question, I study Teach For America (TFA) teachers working in New York City (NYC) between 2012 and 2019. This research brief reports on three key findings. First, I document five-year cumulative retention rates and find that, as expected, retention is lower for TFA teachers (25%) than for other NYC teachers working in similar schools (43%). Second, I estimate within-teacher returns to experience using a teacher fixed effects strategy and find that TFA teachers who continue teaching through year five improve at double the rate of the average NYC teacher. Third, I model the joint relationship between turnover and performance over time and find that the TFA performance advantage is large enough to offset turnover costs. I conclude that the net effect of TFA hiring on student achievement is positive in the short and long run.

VERSION: November 2022

Suggested citation: Lovison, Virginia S.. (2022). The Effects of High-performing, High-turnover Teachers on Long-run Student Achievement: Evidence from Teach For America. (EdWorkingPaper: 22-675). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/9mhe-1w08

The Effects of High-performing, High-turnover Teachers on Long-run Student Achievement: Evidence from Teach For America

Virginia S. Lovison

Brown University

November 2022

Author Note

Correspondence concerning this article should be addressed to Virginia Lovison at virginia_lovison@brown.edu. This research benefited tremendously from feedback from Eric Taylor, Heather Hill, Michela Carlana, and Matthew Kraft. This study would not have been possible without the support and assistance of the New York City Department of Education, Teach For America, and the Center for Educational Policy Research at Harvard University. I have no known conflicts of interest to disclose. All errors are my own.

Abstract

An increasing share of new teachers enter the profession through alternative certification programs. While these programs increase teacher supply in areas facing critical shortages, they also increase instability in local teacher labor markets via high teacher turnover. A fundamental question is what effect these programs have on student achievement over the long run. To address this question, I study Teach For America (TFA) teachers working in New York City (NYC) between 2012 and 2019. This research brief reports on three key findings. First, I document five-year cumulative retention rates and find that, as expected, retention is lower for TFA teachers (25%) than for other NYC teachers working in similar schools (43%). Second, I estimate within-teacher returns to experience using a teacher fixed effects strategy and find that TFA teachers who continue teaching through year five improve at double the rate of the average NYC teacher. Third, I model the joint relationship between turnover and performance over time and find that the TFA performance advantage is large enough to offset turnover costs. I conclude that the net effect of TFA hiring on student achievement is positive in the short and long run.

Introduction

Districts facing staffing shortages often hire teachers from alternative certification programs to fill open teaching positions. An important question in education research is how this hiring strategy impacts students in the long run. A chief concern is that the "revolving door" of inexperienced and minimally prepared teachers could dampen the overall quality of the teaching workforce and hamper student learning over time (Ingersoll, 2004). This concern makes sense because teachers improve rapidly over the first five years of their careers (Rockoff, 2004; Hanushek et al., 2005; Clotfelter et al., 2010; Harris & Sass, 2014; Kane et al., 2008; Papay & Kraft, 2015), and districts miss out on these performance gains when early career teachers leave.

To examine this concern, I study Teach for America (TFA) teachers working in New York City schools. Extensive prior research documents that TFA teachers perform as well or better than their most-likely alternative during their first two years in the classroom (Clark et al., 2015; Chiang et al., 2017; Decker et al. 2004; Glazerman et al., 2006; Hansen et al., 2014; Henry et al. 2014; Kane et al. 2008; Xu et al. 2008; for an exception, see Boyd et al. 2006). However, this research rarely attends to the substantial number of teachers who choose to keep teaching as alumni of the TFA program. Though over one-third (36%) of TFA teachers continue teaching for four years or more (Donaldson & Johnson, 2011), the research base on how these teachers perform after they complete their TFA commitment is thin.[1] This is an important omission because information on the longer-run performance of TFA teachers—not just how TFA teachers perform right out of the gate—is necessary to carefully evaluate the expected downstream impacts of TFA hiring on student achievement.

[1] I am aware of two other papers that examine this research question in detail. The first is Kane, Rockoff, & Staiger (2008) which is further discussed as the paper progresses. The second is contemporaneous work by Sally Hudson using data from North Carolina.

In this research brief, I model the longer-run performance trajectories of TFA teachers working in New York City public schools. Specifically, I estimate within-teacher returns to experience using a standard teacher fixed effects identification strategy (Rockoff 2004). Throughout, I refer to teacher performance as teachers' estimated contribution to student test scores. While an imperfect and incomplete measure of teacher performance, estimates of teacher "value-added" strongly predict meaningful long-run outcomes for students (e.g. Chetty et al., 2014), and are therefore used as a performance proxy throughout. The teacher fixed effect approach isolates variation within individual teachers' own performance over time, such that an individual teacher's performance in year *t* is estimated relative to their own performance in year one.

I find that alumni of the TFA program who chose to keep teaching improve rapidly over the first five years of their career. The average TFA teacher who stays in the classroom for five years or more improves by 0.13 student standard deviations over the early years of their career, which is more than double the rate of early career improvement for non-TFA teachers (0.05 student standard deviations).

I then model the effects of a long-run strategy of hiring TFA teachers relative to a counterfactual hiring strategy of no TFA hiring. This approach closely follows Kane, Rockoff, and Staiger (2008) and adjusts performance expectations to account for TFA teacher turnover. Over the study period, only 25% of TFA teachers kept teaching through year five, compared to 43% of non-TFA teachers working in similar schools. Nevertheless, the results suggest that a policy of continuous TFA hiring boosts steady-state student achievement by an estimated 0.05 standard deviations. So long as exiting TFA teachers are replaced by new TFA recruits, student

achievement will be higher, on average, under a "TFA Hiring" policy condition than a "No TFA Hiring" policy condition.

To be clear, hiring is complex and estimated teacher effects on student achievement should not be the only criterion driving staffing decisions. This brief addresses the specific concern that high turnover among alternatively-certified teachers drives down student achievement over the long run. In the case of TFA, the data does not substantiate this concern given the rapid rate at which TFA teachers improve on-the-job.

Data

The data for this analysis comes from the New York City Department of Education (DOE) and from TFA. The DOE provided standard, administrative student- and teacher-level panel data from 2012-2013 to 2018-2019. The DOE does not maintain data on charter school teachers, so charter school teachers are not included in this analysis. The DOE identified TFA teachers using a roster provided by TFA containing teachers' full name and date of birth, and achieved a match rate of 72%. There are no observable demographic differences between the matched sample and the full population of NYC TFA teachers over the study period.

This analysis relies on two key variables: years of teaching experience and teacher career length. I use the DOE's salary codes to identify teachers with one through six years of total teaching experience. A key advantage of using salary codes is that they provide a more reliable indicator of experience given a teacher who appears in the DOE data for the first time may be new to NYC, but not new to teaching. A key disadvantage is the district top-coded years of experience at six, so I cannot reliably distinguish among teachers with six or more years of experience.

Teacher career length is defined with respect to teachers' tenure with NYC schools. I am not able to observe teachers' employment decisions after they exit the district, though it is plausible that some teachers may leave NYC to teach elsewhere. This does not pose a threat to this paper's central findings regarding the long-run effects of TFA hiring, as whether teachers exit the district or exit the profession is not policy relevant at the district-level.

Sample

I restrict the sample to teachers with available performance data. These are math and English Language Arts teachers working in grades four through eight. Altogether, the sample includes 308 unique TFA teachers and approximately 40,000 non-TFA teachers. On average, TFA teachers work in schools where student performance is 0.36 standard deviations below the city-wide mean. These are also schools serving relatively higher concentrations of students from low-income backgrounds and students from minoritized racial backgrounds.

Figure 1 reports cumulative retention rates for teachers working in schools that hire TFA teachers. The sample risk of leaving the NYC teacher workforce is highest for newly hired teachers, declines steadily for the first four years after a teacher is hired, and then tapers off after year five. This general pattern holds for both TFA and non-TFA teachers, with two exceptions. First, TFA teachers are five percentage points more likely to persist to a second year of teaching (survival probability = 0.87; se=0.01) than non-TFA teachers working in similar schools (survival probability = 0.82; se=0.01). Second, conditional upon teaching for at least two years, non-TFA teachers are retained at higher rates than TFA teachers in years three and beyond. Over the study period, 43% of non-TFA teachers kept teaching through year five, compared to only 25% of TFA teachers. Beyond year 6, retention rates appear to level off at about 33% for non-TFA teachers and 13% for TFA teachers.

Estimation

Following Rockoff (2004), I estimate returns to experience according to Equation 1:

$$A_{ijsgt} = \Delta \sum_{y=1}^{6} \mathbf{1}\{Experience *_{jt} = y\} + \Phi \sum_{y=1}^{6} \mathbf{1}\{Experience *_{jt} = y\} * TFA_j + \alpha A_{ijsg(t-1)} + \delta_j + \pi_t + \Gamma \boldsymbol{X} + \epsilon_{ijtsg}$$

where

$$Experience *_{jt} = Experience_{jt} \quad if \ \ Experience < 6$$
$$= 6 \quad if \ \ Experience \geq 6$$

The outcome, A_{ijsgt}, is the test score for student *i* in subject *s* and grade *g* assigned to teacher *j* in year *t.* Test scores are standardized within grade, subject, and year for all tested students in grades 3-8. The key treatment variables are indicators for teacher experience, and the omitted reference category is first year teacher performance (experience = 1). I interact each treatment variable with an indicator equal to one for TFA teachers and 0 otherwise. The specification includes a teacher fixed effect, δ_j, which isolates variation in performance over time within each individual teacher. The specification also includes students' prior test scores, $A_{ijsg(t-1)}$, which are permitted to vary by subject, grade, and school year; year fixed effects, π_t, which account for average differences in student test scores across time; student demographics including student gender, race, English proficiency, and disability classification; and classroom- and school-level averages of the same student demographic variables (**X**). These student demographic controls help reduce remaining selection bias arising from non-random sorting of students to teachers that is not already captured by test scores. The coefficients of interest are included in the vector, ϕ, and represent the average difference in the returns to experience between TFA and non-TFA teachers for each experience bin.

Because year and experience are collinear for most teachers, I impose a ceiling at six years of experience, such that the year six experience bin includes all teachers with six or more years of experience.[2] This choice overcomes what is known as the age-period-cohort problem but requires the assumption that returns to experience are constant beyond year six. If teachers were to improve beyond year six, the estimated returns to experience would be understated (Papay & Kraft, 2015).

Next, I estimate long-run effects of TFA hiring on student achievement accounting for turnover. These estimates assume no average difference in initial, year one performance for TFA and non-TFA teachers, which is consistent with observational estimates in the study sample. These estimates also pertain to an assumed "steady state" of the world, with no major changes in teacher supply or average teacher performance over time, and a constant turnover rate after year five. The steps below detail the procedure from Kane, Rockoff, & Staiger (2008):

1) Estimate the steady state distribution of teacher experience by TFA participation status. This yields the expected proportion of TFA (non-TFA) teachers with one, two, three, four, or five+ years of experience in any given school year in the steady state.
2) Multiply the expected proportion of TFA (non-TFA) teachers within a given experience bin with average estimated value-added scores for teachers in that experience bin.
3) Sum over all experience bins.
4) Calculate the difference in estimated value-added scores between TFA and non-TFA teachers.

[2] Data limitations necessitate imposing the ceiling at six years. The DOE data does not reliably identify teachers with more than six years of experience.

Results

Figure 2 plots returns to experience for TFA and non-TFA teachers. The results suggest that TFA teachers improve at a faster rate than non-TFA teachers over the first five years of their careers. In particular, between years one and two, TFA teachers improve at twice the rate of non-TFA teachers (p=.069). This TFA performance advantage dissipates in year three as non-TFA teachers appear to catch up, and then re-emerges in years four and beyond. Over the first five years, TFA teachers improve by an estimated 0.13 student standard deviations in teacher value-added whereas non-TFA teachers improve by 0.05 standard deviations, on average.

All estimates normalize initial teacher performance in year one to zero. Consequently, the returns to experience estimates do not, on their own, reveal level differences in TFA and non-TFA teacher performance. However, in practice, cross-sectional estimates of initial teacher performance in the study sample reveal no average difference in first year performance for TFA and non-TFA teachers. Controlling for prior test scores, as well as student-, classroom-, and school-level demographics, the average difference in performance for students assigned to a first year TFA teacher and students assigned to a first year non-TFA teacher was 0.003 standard deviations (se=.011). These results are consistent with experimental estimates of the effect of first-year TFA teachers on student performance (Glazerman et al., 2006). Thus, in this particular case, the performance trajectories depicted in Figure 2 are informative about both level differences in performance between TFA and non-TFA teachers and differences in performance improvement over time. One caveat is that although the returns to experience profile for TFA teachers is increasing over time, I cannot reject that the returns to experience between years one and six are different to the returns to experience between years one and two. For instance, the

TFA confidence interval for returns to experience at year six and beyond [0.053,0.215] overlaps with the confidence interval for the returns to experience in year two [0.027,0.102].

What do these results imply regarding the long-run effects of TFA hiring? I estimate steady state student achievement using the returns to teaching experience (Figure 2) and the proportion of teachers in each experience bin in the steady state (Figure 3). As shown in Figure 3, TFA teachers are about 1.5 times more likely to be in their first year of teaching as non-TFA teachers and only half as likely to hold five years of experience or more. Nevertheless, following the procedure described above, I find that TFA hiring increases steady state student achievement by 0.05 student-level standard deviations. This estimate assumes a constant turnover rate after year five and assumes no average difference in performance for TFA and non-TFA teachers in year one. However, it could be the case that the true counterfactual TFA hire is not the average first year non-TFA teacher, but rather a long-term substitute or an emergency certified teacher. If TFA teachers perform better in year one than the most likely alternative teacher for their position, then the estimates reported here would understate the true effect of TFA hiring on steady state student achievement.

On the other hand, average changes over time in the size or quality of the TFA workforce would threaten the conclusions drawn in this brief. For instance, if changes to the supply of TFA teachers result in conditions where exiting TFA teachers are replaced by non-TFA teachers, the results presented here would no longer hold. Likewise, if the quality of the TFA workforce deteriorated over time due to recruitment challenges, or if the returns to experience reported above are biased upward by the unique set of TFA alumni teachers in the NYC sample, then these predictions would overstate the long-run effects of TFA hiring.

Discussion and Implications

Extensive prior research documents that TFA teachers perform as well or better than their most-likely alternative during their first two years in the classroom. However, this research largely overlooks the now-substantial fraction of TFA teachers who choose to keep teaching as alumni of the TFA program. This paper examined the longer-run performance trajectories of TFA teachers, estimating how TFA teachers improve over time. I find that TFA teachers improve rapidly over the first five years of their career, outperforming non-TFA teachers by 0.08 student-level standard deviations by the end of year five. I also conclude that these performance gains are strong enough to offset the negative costs of high turnover among TFA recruits. In the steady state, continuous TFA hiring boosts student achievement by an estimated 0.05 standard deviations.

Notably, this brief reports returns to experience for TFA teachers that differ from earlier work by Kane, Rockoff, and Staiger (2008) despite the fact that both studies use similar methods and study teachers in NYC. A plausible explanation for why TFA teachers of more recent cohorts improve more quickly than cohorts from the early 2000s is that TFA's workforce and organizational strategy evolved dramatically between the two study periods. Of note, the TFA teacher workforce of the late 1990s and early 2000s skewed young, wealthy, and White, with many graduates from elite universities.[3] Today, this characterization of the TFA workforce no longer holds. As of 2014, one-third of TFA teachers identified as first-generation college graduates, one-half as teachers of color, and one-half as growing up with a low household income (Mead et al., 2015). A consequence of this demographic shift is that recent cohorts of TFA teachers have more in common with the students they teach than TFA teachers from cohorts past. This shared identity may help teachers develop strong relationships with students

[3] See KRS (2008) for demographic characteristics of the NYC TFA teacher sample from 1998-2005: 9% of teachers identified as Black; 9% as Hispanic; median age was 23.

that in turn enhance performance over time. While this study does not test this hypothesis empirically, extensive research documents the positive effects of student-teacher racial congruence on educational outcomes for students of color (Dee, 2004; Egalite et al., 2015; Lindsay & Hart, 2017; Gershenson et al., 2016; Gershenson et al., 2018).

In summary, this study echoes the results of a large body of research documenting positive effects of TFA on student achievement. I contribute to this body of work evidence on how TFA teachers improve over time, with a focus on TFA teachers who complete their two-year commitment and choose to keep teaching. I find that though TFA teachers turnover at relatively high rates, they also improve rapidly over the years they choose to teach. Overall, the performance of the TFA workforce is strong enough to offset turnover. A ripe area for future research is to examine whether improvements in the performance of the TFA workforce since the early 2000s is a direct result of TFA's efforts to recruit a teaching force that more closely resembles the students it serves.

References

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement. Education Finance and Policy, 1(2), 176-216.

Chiang, H., Clark, M., & McConnell, S. (2017). Supplying Disadvantaged Schools with Effective Teachers: Experimental Evidence on Secondary Math Teachers from Teach For America. Journal of Policy Analysis and Management, 36(1), 97-125.

Clark, M. A., Isenberg, E., Liu, A. Y., Makowsky, L., & Zukiewicz, M. (2015). Impacts of the Teach For America investing in innovation scale-up. *Princeton, NJ: Mathematica Policy Research*.

Clotfelter, C., Ladd, H. & and Vigdor, J. (2010). Teacher Credentials and Student Achievement in High School. The Journal of Human Resources 45, no. 3 (2010): 655–81.

Decker, P. T., Mayer, D. P., & Glazerman, S. (2004). *The effects of Teach for America on students: Findings from a national evaluation*. University of Wisconsin--Madison, Institute for Research on Poverty.

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.

Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: The impacts of Teach for America on student achievement and other outcomes. Journal of Policy Analysis and Management, 25(1), 75-96.

Hansen, M., Backes, B., Brady, V., & Xu, Z. (2014). Examining spillover effects from Teach For America corps members in Miami-Dade County Public Schools. Washington, DC: the CALDER Center at the American Institutes for Research.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). The Market for Teacher Quality. NBER Working Paper No. 11154. *National Bureau of Economic Research*.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of public economics*, *95*(7-8), 798-812.

Henry, G. T., Bastian, K. C., Fortner, C. K., Kershaw, D. C., Purtell, K. M., Thompson, C. L., & Zulli, R. A. (2014). Teacher Preparation Policies and Their Effects on Student Achievement. Education Finance and Policy, 9(3), 1-40.

Ingersoll, R. M. (2004). Why do high-poverty schools have difficulty staffing their classrooms with qualified teachers?.

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

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*(4), 476-500.

Mead, S., Chuong, C., & Goodson, C. (2015). *Exponential Growth, Unexpected Challenges. How Teach for America Grew in Scale and Impact*. Bellwether Education.

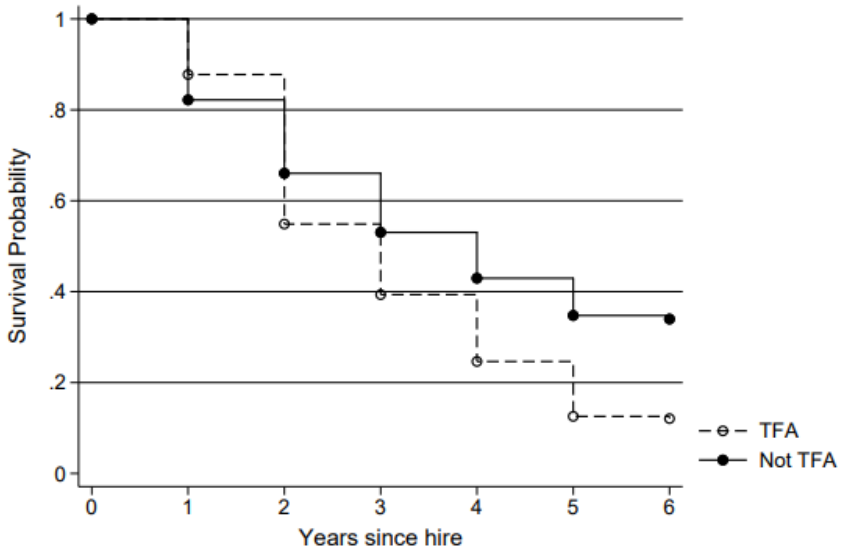
Papay, J. P., & Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics*, *130*, 105-119.

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

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.

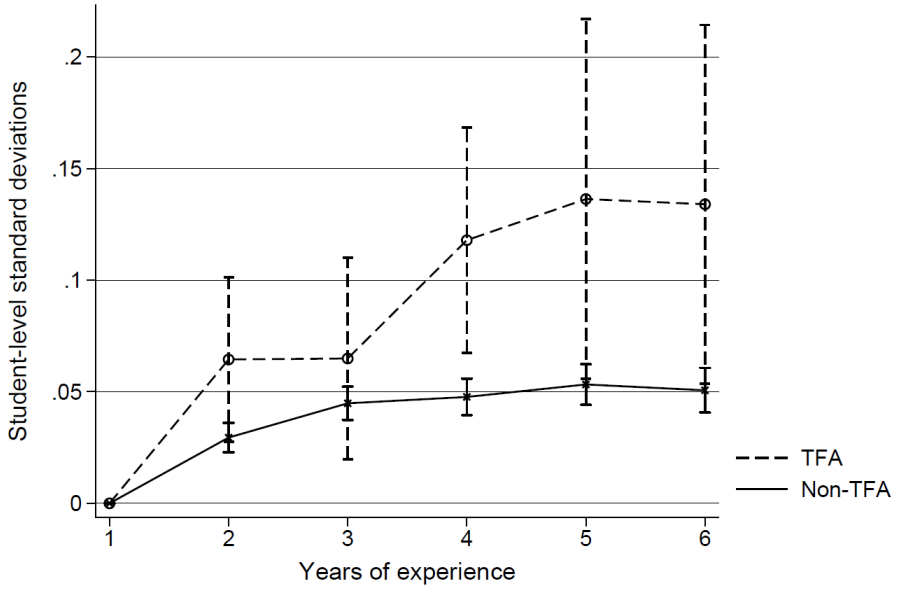
Figures

Figure 1. Cumulative retention rates



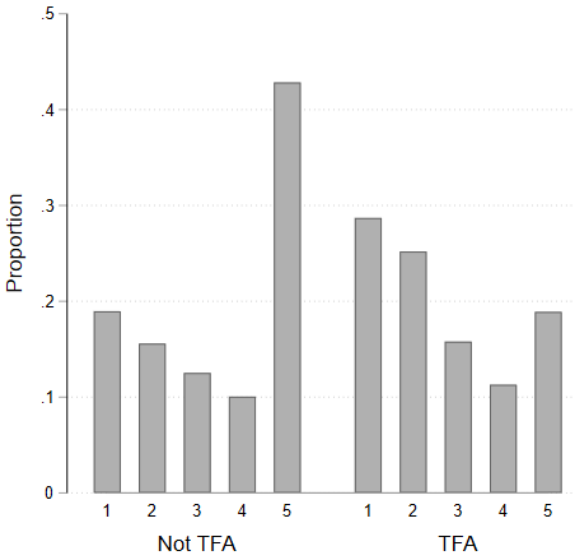
Note. This figure plots survival probabilities for TFA and non-TFA teachers. The sample includes all newly hired teachers working in schools that hire TFA teachers. I construct a teacher-level dataset with the total years of experience for each teacher and an indicator variable equal to one if a teacher left the NYC teacher workforce. The plotted estimates represent the sample retention probabilities for teachers in the risk set at each time interval (i.e. each interval excludes those who exited the teaching workforce in a prior year).

Figure 2. Returns to teaching experience



Note. This figure plots estimates from a single linear regression. The dataset is organized at the student-subject-year level. The dependent variable is a student's test score in math or English Language Arts. Test scores are standardized within grade, subject, and year for all students in grades 3-8 with non-missing test scores. The key treatment variables are indicators for teacher experience for years 1, 2, 3, 4, 5, and 6+. The omitted category is defined as teachers with one year of experience. Each treatment variable is interacted with an indicator equal to one for TFA teachers. The specification also includes students' prior test scores, which are permitted to vary by subject, grade, and school year; teacher fixed effects; year fixed effects; student demographics including student gender, race, English proficiency, and disability classification; and classroom- and school-level averages of the same student demographic variables. The performance trajectory is denoted by the dashed line for TFA teachers and the solid line for non-TFA teachers. The vertical lines depict the 95 percent confidence interval for each estimate. Standard errors are clustered at the teacher level.

Figure 3. Steady state distribution of teacher experience, by TFA participation status



Note. This figure plots the steady state proportion of teachers in each experience bin for TFA and non-TFA teachers. Experience bin 5 contains all teachers with five or more years of experience.

Appendix

Table A1: Returns to teaching experience

2 yrs exp	0.029
(relative to a first year teacher)	(0.003)
3 yrs exp	0.045
	(0.004)
4 yrs exp	0.048
	(0.004)
5 yrs exp	0.053
	(0.005)
6 yrs exp	0.051
	(0.005)
TFA * 2 yrs exp	0.035
	(0.019)
TFA * 3 yrs exp	0.020
	(0.023)
TFA * 4 yrs exp	0.070
	(0.026)
TFA * 5 yrs exp	0.083
	(0.042)
TFA * 6 yrs exp +	0.083
	(0.041)
N of students	5064939
N of teachers	32917

Note. This table presents estimates from a single linear regression. Standard errors are provided in parentheses. The dataset is organized at the student-subject-year level. The dependent variable is a student's test score in math or English Language Arts. Test scores are standardized within grade, subject, and year for all students in grades 3-8 with non-missing test scores. The key treatment variables are indicators for teacher experience for years 1, 2, 3, 4, 5, and 6+. The omitted category is defined as teachers with one year of experience. Each treatment variable is interacted with an indicator equal to one for TFA teachers. The specification also includes students' prior test scores, which are permitted to vary by subject, grade, and school year; teacher fixed effects; year fixed effects; student demographics including student gender, race, English proficiency, and disability classification; and classroom- and school-level averages of the same student demographic variables. Standard errors are clustered at the teacher level.