

IZA DP No. 7296

**The Effect of Teach for America on the Distribution of  
Student Achievement in Primary School:  
Evidence from a Randomized Experiment**

Heather Antecol  
Ozkan Eren  
Serkan Ozbeklik

March 2013

# **The Effect of Teach for America on the Distribution of Student Achievement in Primary School: Evidence from a Randomized Experiment**

**Heather Antecol**

*Claremont McKenna College and IZA*

**Ozkan Eren**

*Louisiana State University*

**Serkan Ozbeklik**

*Claremont McKenna College*

Discussion Paper No. 7296  
March 2013

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **The Effect of Teach for America on the Distribution of Student Achievement in Primary School: Evidence from a Randomized Experiment**

Using data from a randomized experiment and fixed effect quantile regression (FEQR), we look at the effects of having a TFA teacher on test scores across the entire achievement distribution of primary school students in disadvantaged neighborhoods. While we find that TFA teachers neither help nor hurt students in terms of reading test scores, we find positive and statistically significant effects of TFA over the entire math achievement distribution for the full sample and the effects are fairly uniform. We find a similar effect of TFA across the math test score distribution irrespective of student gender, although the FEQR estimates for female students are two to three times larger than for male students. In addition, we find that there is significant heterogeneity in the effects of TFA for Hispanic and black students and for students taught by novice teachers. Finally, we find that the effect of TFA is homogeneous across the math achievement distribution irrespective of certification type. Taken together, these patterns suggest that allowing highly qualified recent college graduates and mid-career professionals, who in the absence of TFA would not have taught in these disadvantaged neighborhoods, should have a positive influence not just on students at the top of the math achievement distribution but across the entire math achievement distribution.

JEL Classification: C21, I21, I28, J24

Keywords: student achievement, random assignment, fixed effect quantile regression

Corresponding author:

Heather Antecol,  
The Robert Day School of Economics and Finance  
Claremont McKenna College  
500 E. Ninth St.  
Claremont, CA 91711  
USA  
E-mail: [hantecol@cmc.edu](mailto:hantecol@cmc.edu)

## 1. Introduction

“...we’ve got an obligation to lift up every child in every school in this country, especially those who are starting out furthest behind...The whole premise of Race to the Top is that teachers are the single most important factor in a child’s education from the moment they step into the classroom.... So I am 110 percent behind our teachers. But all I’m asking in return – as a President, as a parent, and as a citizen -- is some measure of accountability.” President Obama, July 29, 2010, the Urban League’s 100<sup>th</sup> Anniversary Convention.<sup>1</sup>

The United States continues to lag behind many other developed countries in terms of student test score outcomes (Fleischman et al. (2010)). How one can improve student outcomes is subject to great debate. The Obama Administration is continuing to move away from a school accountability system as evidenced by the No Child Left Behind (NCLB) Waivers<sup>2</sup> and is moving towards a teacher accountability system as evidenced by the Obama Administration’s continued commitment to the Race to the Top (RTTT) Program.

The RTTT Program places an extremely high weight (i.e., roughly a 1/3 of the 500 point allotment) on criteria that enhances “great teachers and leaders” including alternative routes to certification and linking teacher/principal effectiveness to student performance. RTTT also emphasizes, although to a lesser degree (i.e., roughly a 1/10 of the 500 point allotment), “turning around the lowest-achieving schools”.<sup>3</sup> These goals of the Program tend to go hand in hand given teachers with alternative certifications tend to teach in disadvantage neighborhoods with lower achieving students (see Kane et al. 2008 and Darling-Hammond et al. 2005 and references therein).

<sup>1</sup><http://www.whitehouse.gov/the-press-office/remarks-president-education-reform-national-urban-league-centennial-conference>.

<sup>2</sup> NCLB Waivers have been issued to the following states as of January 2013: Arizona, Arkansas, Colorado, Connecticut, DC, Delaware, Florida, Georgia, Idaho, Indiana, Kansas, Kentucky, Louisiana, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Nevada, NJ, NM, NY, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, South Dakota, Tennessee, Utah, Virginia, Washington, and Wisconsin.

<sup>3</sup> The RTTT Program also awards points for “state success factors”, “standards and assessment”, “data systems to support instruction”, “general selection criteria”, and STEM (U.S. Department of Education, 2009).

Given the goals of RTTT, it is important to understand the role of alternative routes to certification. Should only traditionally certified teachers be allowed to teach or should alternative routes to certification be allowed? Proponents of certification argue that approved education programs ensure that teachers will be of high quality. Opponents, on the other hand, argue certification itself does not ensure quality as being a “good” teacher is not only a function of being trained on how to teach but it is also a function of overall academic ability and substantial subject matter knowledge. Moreover, opponents argue that certification imposes barriers to entry that differ by race/ethnicity which is particularly problematic for highly disadvantaged neighborhoods.<sup>4</sup>

One of the largest scale alternative routes to certification is the Teach for America (TFA) Program.<sup>5</sup> Teach for America (TFA) is a non-profit organization that recruits outstanding recent college graduates and mid-career professionals to teach in schools in highly disadvantaged neighborhoods throughout the United States for at least a two-year period. Although each TFA corps member goes through a very competitive selection process and receives a short but very intense training program including significant support following training<sup>6</sup>, critics argue that TFA teachers are still ill-prepared to teach in the challenging environments due to the lack of

<sup>4</sup> See Lackzko-Kerr and Berliner (2002) and Darling et al. (2005) for a more detailed discussion of the views of proponents and opponents of teacher certification.

<sup>5</sup> Other types of alternative routes to certification include, but are not limited to, emergency (issued for a specified period to persons with insufficient teacher preparation), temporary (given to persons who need additional preparatory training, such as coursework and/or student teaching, in order to satisfy requirements for full certification), and initial (issued to candidates after satisfying all requirements except the completion of a probationary period) certificates (Decker et al. 2004).

<sup>6</sup> To be eligible for TFA, applicants must have a minimum GPA of 2.5 at the time of application (and when they graduate) and candidates must meet several criteria, including but not limited to, possession of personal responsibility, motivational and organizational ability, critical thinking and commitment to the TFA mission. Once selected, TFA corps members must participate in a five-week summer institute designed to help them to become an effective teacher in a low-income community. Specifically, the summer institute includes six formal education courses, teaching for a class of summer school students as well as several meetings and workshops on teaching methods. Corp members also receive ongoing support from TFA staff/faculty during their TFA experience.

specialized education, training, and experience (Lackzo-Kerr and Berliner 2002, Pilcher and Steele 2005, Darling-Hammond et al. 2005).<sup>7</sup> With that said, since its inception in 1989, approximately 24,000 TFA Corps Members have taught more than 3 million students in 38 urban and rural areas. Between 2000 and 2009 the number of applications for TFA skyrocketed to 35,000 from 4,068 and the number of new corps members recruited each year grew from 868 to 4,100.

The evidence on teacher effectiveness (as measured by student test scores) based on the TFA program is mixed at best.<sup>8</sup> While some studies find that TFA teachers are less effective than traditionally certified teachers and perform as well as other uncertified teachers (see for example, Darling-Hammond et al. 2005 and Laczko-Kerr and Berliner 2002), others find that TFA teachers perform as well or better than both traditionally certified and other uncertified teachers (see for example, Decker et al. 2004, Kane et al. 2008, and Raymond et al. 2001). In part, the results vary across studies due to differences in their methodological approach and their data. Only one study, to the best of our knowledge, relies on experimental data (Decker et al. 2004) which avoids the issue of non-random assignment of teachers most of the other research could not account for. Specifically, if high achieving and better motivated students are less

<sup>7</sup> Critics also argue that TFA teachers are more likely to exit the teaching profession than their non-TFA counterparts. This seems less plausible in light of the fact that TFA tends to place their recruits in schools with low achieving students (Decker et al. 2004) and schools with low achieving students tend to have higher teacher exit rates (Hanushek et al. 2004).

<sup>8</sup> The effect of traditionally certified teachers relative to teachers with other alternative certifications on student achievement has also received considerable attention in the literature (see for example, Goldhaber and Brewer 2000; 2001 and Jepsen and Rivkin 2009). In a recent study by Constantine et al. (2009), which is based on random assignment, finds little evidence to support the notion that teachers with less selective alternative certifications (excluding TFA) adversely affects student outcomes in disadvantaged schools (although somewhat less disadvantaged than schools in the TFA sample) relative to their traditionally certified counterparts.

likely to be assigned to TFA teachers, then the effect of having a TFA teacher on achievement will be understated unless this nonrandom selection in the observational data is taken into account in the estimation.

All of the existing literature tends to focus on the impact of TFA on the mean achievement outcomes of students. Focusing solely on the mean impact however may mask meaningful and policy relevant heterogeneity across the distribution. As such, the program evaluation literature has turned its focus on the effects at the distributional level. Specifically, there is evidence of considerable heterogeneity in the program response for job training (Eren and Ozbeklik 2012, Abadie et al. 2002, and Heckman et al. 1997), welfare (Bitler et al. 2006, 2008 and Friedlander and Robins 1997), re-employment bonuses (Koenker and Biliias 2001), and education reforms (Jackson and Page 2013). Considering the significant heterogeneity in the impacts of these programs, as well as the Obama's administration's continued support of employing alternative routes to certification and improving schools in disadvantaged neighborhoods, it is perhaps surprising that no work to date has examined the heterogeneity in the impacts of the TFA program on student test score outcomes.

In this paper, we attempt to fill this gap in the literature using unique data from a well-executed randomized experiment that was conducted to evaluate the effectiveness of the TFA program. Specifically, we look at the effect of having a TFA teacher across the entire achievement distribution (i.e. reading and math test scores) using a fixed effects quantile regression (FEQR) approach recently introduced in Canay (2011). Focusing first on reading test scores, we find that TFA teachers neither help nor hurt students and this is true throughout the test score distribution irrespective of the sample under consideration.

Turning to math test scores, we find positive and statistically significant effects of TFA over the entire achievement distribution for the full sample and the effects are fairly uniform. These overall findings however mask some important differences across subgroups. While we find a similar effect of TFA across the math test score distribution irrespective of student gender, the FEQR estimates for female students are two to three times larger than for male students. In addition, we find that there is significant heterogeneity in the effects of TFA for Hispanic and black students and for students taught by novice teachers. Finally, we find that irrespective of certification type, TFA teachers have a positive effect on math test score outcomes across the math achievement distribution. Taken together, these patterns suggest that allowing highly qualified recent college graduates and mid-career professionals, who in the absence of TFA would not have taught in these disadvantaged neighborhoods, should have a positive influence not just on students at the top of the math achievement distribution but across the entire math achievement distribution.

The remainder of the paper is organized as follows. Section 2 describes the data. Sections 3 and 4 discuss the estimation strategy and the results, respectively. Conclusions are presented in Section 5.

## **2. Data**

We use data from the Mathematica Policy Research, Inc (MPR) National Evaluation of Teach for America (NETFA) Public Use File. NETFA is a randomized study of primary school students in six regions in the United States between 2001 and 2003. The pilot study was conducted in Baltimore in the 2001-2002 academic year followed by full-scale evaluations in Chicago, Los Angeles, Houston, New Orleans and the Mississippi Delta in the next academic

year (2002-2003). Each region had one school district participating in the experiment with the exception of Mississippi Delta which had two school districts that participated in the experiment. Within each school district, schools were selected to reflect where TFA placed teachers at the time of the study and only schools that had both TFA teachers and control teachers in the same grade were considered eligible for the study. The final sample consisted of 17 schools, 100 teachers, and more than 1900 students in grades 1 through 5.

The students in NETFA were randomly assigned to two types of classes, TFA and control group classrooms, before the start of the academic year. Therefore the randomization is done at the block level such that each block represents classrooms in the same grade level in any given school. Furthermore, throughout the year roster checks were performed to enforce the original assignment. After the random assignment and before the start of academic year, the students were given math and reading tests based on the grade they had completed in the previous academic year (pre-treatment test scores); then at the end of the academic year in which the study was conducted the students re-took math and reading tests based on the grade they just completed (post-treatment test scores).

The post-treatment Normal Curve Equivalent (NCE) points in math and reading tests constitute our dependent variables. In a randomized experiment setting, it is a well-known fact that controlling for the baseline characteristics does not affect the consistency of the treatment effect estimate; however, it helps increase efficiency (Frölich and Melly, 2010). To this end, our regression specifications include pre-treatment subject specific test scores, gender, race/ethnicity, eligibility for free or reduced-price lunch, and the percentage of students in the classroom who were not in research sample (See Appendix Table 1 for variable definitions). Finally, given the

clustered nature of the NETFA, it is important to control for school-grade fixed effects. As such, we also control for block fixed effects in all our specifications.

The first four columns of Table 1 report summary statistics for the TFA and control group, respectively. Under the assumption that the random experiment is implemented correctly, the baseline characteristics of students must be similar in TFA and control group classrooms conditional on block fixed effects. To test this, as in Krueger and Whitmore (2001), we run a regression of a TFA indicator variable on each baseline characteristic. The dependent variable takes a value of one if the student is taught by a TFA teacher and zero if the student is taught by a control teacher. The final column of Table 1 reports the regression coefficients from this exercise. None of the coefficient estimates are statistically significant at conventional levels.

One other concern about random experiments is the potential contamination due to attrition. In their report on TFA, Decker et al. (2004) show the overall response rate for NETFA to be very high; around 90 percent. With that said, however, we still lose 10 (12) percent of the initial math (reading) sample because of missing test scores and/or students moving out of the school district. If the sample attrition were to differ between the TFA and control group samples, this could then potentially contaminate the TFA impact. To examine if the nonrandom attrition differs between the TFA and control groups, we run a regression of a non-response indicator variable (taking a value of one if the post-treatment test score is missing and zero otherwise) on a TFA indicator variable along with the baseline characteristics and block fixed effects. The coefficients on TFA are 0.028 (0.022) and 0.007 (0.012) for math and reading test score samples, respectively, and are not statistically different from zero at conventional levels.

## ***2.1 TFA and Control Group Teachers***

Numerous studies have documented the tendency of qualified teachers to move towards schools that serve relatively well-off children even though salaries are often similar in such schools (see for example, Clotfelter et al. 2008, Hanushek et al. 2004, Scafidi et al. 2007). By design, the NETFA Public Use File was chosen to be broadly representative of the schools where TFA teachers taught at the time of the evaluation and these schools are disadvantaged schools. It is therefore important to understand who normally teaches in these schools. Following Decker et al. (2004) we examine three different categories of teachers from our effective sample: (i) control group teachers (irrespective of teaching experience), (ii) novice control group teachers (three or fewer years of teaching experience) and (iii) TFA teachers. We include the second category as this group closely represents the set of teachers who would have been hired by schools had TFA not been available.

Table 2 reports several teacher characteristics including demographics, credentials, teaching methods, and teaching beliefs. Most control group teachers are female and non-white in our sample. This holds for the entire sample of control group teachers, as well as for novice control teachers. Relative to the control group teachers, TFA teachers, on the other hand, are less likely to be female, are younger and a majority of them are white. Relative to the entire pool of control group teachers, a smaller percentage of TFA teachers hold traditional teacher certification, while this fraction is larger compared to novice control teachers. With respect to in class teaching practices, we do not observe significant gaps among control group and TFA teachers. The selected teaching belief questionnaires, on the other hand, indicate remarkable differences. For instance, 62% of control group teachers (52% of novice control teachers) place a major emphasis on memorizing facts, rules and steps, compared to only 26% of TFA teachers,

while 57% of control group teachers (47% of novice control teachers) place a major emphasis on getting the right answer, compared to only 9% of TFA teachers. Finally, TFA teachers are more likely than control group teachers to believe that computational speed and accuracy is important. These differences in teaching styles appear to be consistent with the types of training TFA recruits receive as well as the types of recruits attracted to and recruited for the TFA program.<sup>9</sup>

### **3. Empirical Methodology**

#### ***3.1 Mean Approach***

To initially examine the data, we utilize a standard regression approach and thereby focus on the conditional mean. Specifically, we estimate a linear regression model of the form

$$TS_{ijb} = \alpha TFA_{ijb} + X_{ijb}\beta + \eta_b + \varepsilon_{ijb} \quad (1)$$

where  $TS$  is the post-treatment test score for student  $i$  in class  $j$  and block  $b$ ,  $TFA$  is a binary indicator variable taking the value of one if the class of the student is taught by a TFA teacher and zero if the class of the student is taught by a control teacher and  $X$  is a vector of covariates (student gender, student race/ethnicity, eligibility for free or reduced lunch, and percentage of students in the classroom who were not in the research sample). Finally,  $\eta$  represents block fixed effects, and  $\varepsilon$  is an error term with the standard properties.

<sup>9</sup> For an interesting discussion of teacher traits, see the Atlantic article “What Makes a Great Teacher?” by Ripley (January/February Issue, 2010).

### 3.2 Fixed Effect Quantile Regression (FEQR) Approach

Focusing on the mean however may mask meaningful and policy relevant heterogeneity across the distribution. The common practice in the literature to examine such heterogeneity is to utilize the traditional quantile regression (QR) approach.<sup>10</sup> The block fixed effects in equation (1), however, do not allow us to use the QR approach. This follows from the fact that standard demeaning techniques rely on the linearity of the expectation operators, which is not the case for conditional quantiles. Recently, Canay (2011) proposed a simple transformation of the data which allows one to make inferences throughout the distribution in the presence of fixed effects.<sup>11</sup> Specifically, consider the following structural equation

$$TS_{ijb} = Z_{ijb}\theta + \eta_b + \varepsilon_{ijb} \quad (2)$$

where  $TS$ ,  $\eta$ , and  $\varepsilon$  are as defined previously and  $Z$  is a vector of variables including  $TFA$  and  $X$ . As noted, a straightforward QR estimation of equation (2) is not plausible because of the block fixed effects. Canay (2011) developed a two-step procedure that would produce consistent QR estimates  $\hat{\theta}(\tau)$  for  $0 < \tau < 1$ . The method relies on the following two crucial assumptions:

1.  $\eta$  has a pure location shift effect.
2.  $\varepsilon \perp (Z, \eta)$  and  $\varepsilon \sim U(0, 1)$ .

<sup>10</sup> The QR method was initially introduced by Koenker and Bassett (1978) as a possible efficient alternative to OLS in linear models. Among many others, Buchinsky (1994, 1998) and Powell (1986) extended the use of QR to obtain information about the effects of exogenous explanatory variables on the dependent variable at different parts of the distribution.

<sup>11</sup> There are several other methods that were proposed in the literature to account for fixed effects in a QR setting (see for example, Rosen 2009, Abrevaya and Dahl 2008, Koenker 2004). However, Canay (2011) uses Monte Carlo simulations to show that under most circumstances his method outperforms these alternatives. See his paper for a more detailed discussion.

The first assumption states that observed and unobserved variables captured by the fixed effects must have coefficients that are constant over  $\tau$  and the second assumption requires independence of the error term. These assumptions are likely to hold in our context given the randomized nature of our data set at the school and grade level (blocks).

In practice, under these assumptions the estimation strategy works as follows:

(i) Run an OLS regression of  $TS$  on  $Z$ , obtain  $\hat{\theta}$  and define

$$\hat{\eta}_b = E_b[TS_{ijb} - Z_{ijb}\hat{\theta}].$$

(ii) Let  $\widehat{TS}_{ijb} = TS_{ijb} - \hat{\eta}_b$  and run a standard QR of  $\widehat{TS}$  on  $Z$  to obtain fixed-effects quantile regression (FEQR) parameters,  $\hat{\theta}(\tau)$  for  $0 < \tau < 1$ .

#### 4. Results

The FEQR estimates for math test scores are displayed in Panels A through H of Figure 1 along with the mean effect plotted as a horizontal dashed line. The 0-line is provided for reference. To facilitate comparison along the distribution, we also provide the mean (first column) and the selected FEQR (columns 2-6) estimates in Table 3. The heteroskedasticity-robust standard errors clustered at the block level are reported for OLS estimates, while the standard errors based on 500 bootstrapped cluster replications are given for FEQR estimates (for each bootstrap replication, we re-estimate steps (i) and (ii) of the estimation procedure described above). We do not discuss the reading test score results as the TFA impact on reading scores are usually imprecisely estimated and are close to zero in magnitude over the entire distribution (see Appendix Figure 1). This is true for the full sample analysis and the subgroups (i.e., gender, race, teacher's experience, and teacher's certification) we analyze. In other words, TFA teachers

(relative to control group teachers) seem to neither help nor hurt students in terms of their reading test scores over the entire reading achievement distribution.

#### ***4.1 Full Sample Analysis***

Panel A of Table 3 and Figure 1 indicate a statistically significant TFA impact of around 2.9 points on math test scores, on average. Taking the pre-treatment math test scores from Table 1 as our benchmark, a 2.9 point increase corresponds to an improvement of 0.15 percent of a standard deviation in math test scores. This is consistent with Decker et al. (2004). Turning back to the distributional estimates, we observe a statistically significant impact of TFA over the entire math test score distribution and the effect is fairly uniform in the range of 2 to 3 points for the entire distribution. These results suggest that allowing highly qualified teachers, who in the absence of TFA would not have taught in these disadvantaged neighborhoods, should have a positive influence not just on students at the top of the achievement distribution but across the entire math test score distribution. These findings are particularly important in disadvantaged neighborhoods where students in general tend to have lower test scores than the national average (see Antecol et al. 2013 and references therein).

Focusing on the full sample however may be masking some important differences across subgroups (i.e., gender, race, teacher's experience, and teacher's certification). The remainder of the paper provides a detailed analysis by subgroups.

#### ***4.2 Subgroup Analysis***

**Gender** Our first subgroup results pertain to student gender. Panel B of Table 3 and Figure 1 present the mean and FEQR estimates for male students, while Panel C of Table 3 and Figure 1

present the results for female students. Focusing on male students, the FEQR estimates falls short of statistical significance up until roughly the 65th quantile but are statistically significant afterwards. Precision set aside and taken at face value, the estimates reveal a pattern similar to that of the full sample; the coefficients do not exhibit much heterogeneity over the distribution and usually evolve around the mean effect of 2.1 points.

Turning to female students, the FEQR estimates are once again fairly uniform throughout the distribution. Unlike male students, however, the estimates are all precisely estimated between the 10th and 90th quantiles and are much larger in magnitude. For instance, the effect of being taught by a TFA teacher is 4.2 points for female students at the 50th quantile, while it is only 1.3 points for male students at the same quantile and the difference between these two coefficient estimates is statistically significant at conventional levels.

Given female students tend to lag behind male students in math test scores and this gender gap tends to widen at higher levels of education (see Bedard and Cho 2010 and references therein), the fact that TFA has a larger effect for female student math test scores is encouraging. Specifically, it should help to narrow the gender gap in math achievement and may increase the probability that female students in disadvantaged neighborhoods pursue higher education in math related subjects.

**Race/Ethnicity** The next set of results pertains to student race/ethnicity. Given the focus of TFA on highly disadvantaged neighborhoods, the NETFA study includes only a very limited number of non-Hispanic white students (48 non-Hispanic white students in the math sample). As such, we focus our analysis on black and Hispanic students only. Looking at Panel D of Table 3 and Figure 1, the FEQR estimates reveal large (around 3-3.5 points) and mostly significant effects of

TFA until the 20th quantile for black students. However, between the 20th and 50th quantiles the effect is around 2 points and mostly statistically insignificant. Finally, beginning with the 60th quantile, the FEQR estimates reach 4 points and are statistically significant, and, with an exception at the 80th quantile, remain roughly at that value afterwards. While these results provide suggestive evidence in favor of a slight U-shape in terms of the effects of TFA teachers on black students' math test scores, it still appears to be the case that TFA teachers (relative to control teachers) generally have a positive influence not just on black students at the top of the achievement distribution but across the entire math test score distribution (although the impact is imprecisely estimated at certain points along the distribution).

Turning to the results for Hispanic students (Panel E of Table 3 and Figure 1), we observe non-negligible heterogeneity in the impacts of being taught by a TFA teacher. The effects are precisely estimated up to the 45th quantile and the FEQR estimates exhibit a sharp decreasing trend as we move along the quantile index. For instance, the FEQR estimates indicate a statistically significant impact of roughly 5.0 points at the 10th quantile, while the effect turns out to be -2.2 points at the 90th quantile, though the impact at the 90<sup>th</sup> quantile is imprecisely estimated. The range of point estimates is quite large [4.9, -2.25] and is larger than the 90% confidence interval range constructed around the mean effect.

It is unclear why the benefits from TFA cease to exist at the higher end of the achievement distribution for Hispanic students. One possibility is the small sample of Hispanic students is particularly thin at the upper end of the distribution. Another is that Hispanic students at the higher end of the achievement distribution may be more sensitive to having teachers of the same race/ethnicity and TFA teachers are more likely to be non-Hispanic white than their non-TFA counterparts.

**Teacher Experience** One aspect of teacher quality that receives considerable attention in the education literature is the effect of teacher experience on student achievement. As noted in Decker et al. (2004), a better test of TFA may be to compare TFA teachers to less experienced teachers, not against a mixed group of teachers that contains some highly experienced teachers. This type of comparison is particularly important given the evidence from recent studies showing that a teacher's experience influences student achievement (see for example, Aaronson et al. 2007, Clotfelter et al. 2006, Rockoff 2004) and that a teacher tends to perform worse in their first two years in the classroom (see for example, Rivkin et al. 2005). Given this, we examine the impacts of TFA relative to novice control teachers (i.e., control teachers with three or fewer years of experience).<sup>12</sup>

Panel F of Table 3 and Figure 1 display the TFA impact. The FEQR estimates are statically significant between the 10th and 65th quantiles. In other regions of the distribution, the 90% confidence interval for the FEQR estimates include zero. Moreover, the TFA effect is monotonically decreasing over the math test score distribution. For instance, the impact of being taught by a TFA teacher as opposed to a novice control teacher at the 10th quantile is 6.9 points, which implies a 0.37 standard deviation improvement in math test scores. The impact at the 90<sup>th</sup> quantile, on the other hand, is only 0.1 points. The range of point estimates is quite large [6.9, 0.1].

Therefore if the counterfactual is that students in disadvantaged neighborhoods will be taught by teachers with lower levels of experience (for evidence of this see Hanushek et al.

<sup>12</sup> We also compare TFA teachers to experienced control group teachers (greater than 3 years of experience). The FEQR estimates are predominantly insignificant which prevents us from making any firm conclusions. These results are available upon request.

2004), then allowing TFA teachers to teach in these schools has the ability to have an extremely large impact, particularly at the bottom of the achievement distribution.

**Teacher Certification** One other aspect of teacher quality that receives considerable attention in the education literature is the effect of teacher certification on student achievement (see for example, Goldhaber and Brewer 2000; 2001 and Jepsen and Rivkin 2009). As discussed above, unlike many control group teachers, TFA teachers were primarily non-education majors in college and they were less likely to possess traditional teacher certification when they initially entered teaching. To examine the role of traditional teacher certification, our final set of subgroup results pertains to teacher certification status. Panel G of Table 3 and Figure 1 compares TFA teachers with traditionally certified teachers, while Panel H compares TFA teachers with uncertified teachers. It is hard to make a distinction between these two sets of results. The FEQR regression estimates in both panels are fairly uniform throughout the math test score distribution and evolve predominantly around the mean impact. The coefficient estimates indicate generally statistically significant effects in the range of 3 to 4.5 points over the entire distribution in both panels. These patterns seem to further support the notion that the TFA program appears to do a good job at selecting and training their recruits to teach in extremely disadvantaged neighborhoods.

## 5. Conclusions

Now more than ever understanding the effects of alternative routes to teacher certification is extremely important in light of the Obama Administration's continued emphasis on the Race to the Top (RTTT) Program which places a high value on employing alternative routes to certification and improving schools in disadvantaged neighborhoods. This paper attempts to contribute to the discussion on teacher effectiveness in education policy by thoroughly examining one of the largest scale alternative routes to certification, the Teach for America (TFA) program. The TFA program is particularly relevant for this question as by design TFA places outstanding recent college graduates and mid-career professionals in disadvantaged schools throughout the United States.

While a number of studies have examined the effect of the TFA program on student achievement outcomes (Kane et al. 2008, Darling-Hammond et al. 2005, Decker et al. 2004, Laczko-Kerr and Berliner 2002, and Raymond et al. 2001), these studies focus on the impact of TFA at the mean achievement level of students. Focusing solely on the mean impact however may mask meaningful and policy relevant heterogeneity across the distribution. As such, we look at the effect of having a TFA teacher relative to a control teacher across the entire achievement distribution (i.e., reading and math test scores) using a fixed effects quantile regression (FEQR) approach recently introduced in Canay (2011). We first analyze reading test scores and find that TFA teachers neither help nor hurt students and this is true over the entire reading test score distribution irrespective of the sample under consideration.

Focusing now on the math test score results, we find positive and statistically significant effects of TFA on math test scores over the entire distribution for the full sample and the effects are fairly uniform. These results suggest that allowing highly qualified teachers, who in the

absence of TFA would not have taught in these disadvantaged neighborhoods, should have a positive influence not just on students at the top of the achievement distribution but across the entire math test score distribution.

While these overall findings are particularly important as they provide suggestive empirical support for the emphasis RTTT places on improving teacher effectiveness through alternative routes to certification and improving schools in disadvantaged neighborhoods, they mask some important differences across subgroups. Specifically, we find that the effect of TFA is similar across the math achievement distribution irrespective of student gender, although the FEQR estimates for female students are two to three times larger than for male students. Understanding the mechanisms for why the impact of TFA teachers is substantially larger for female students than for male students requires further investigation which is beyond the scope of this study. We posit however that this may stem from the fact that male students tend to outperform female students in math (although it is unclear if this is true across the entire achievement distribution) and therefore have less room for improvement.<sup>13</sup>

We also find tentative evidence for a positive and slightly U-shaped impact of TFA on the math test scores of black students although the effect tends to be imprecisely estimated between the 20th quantile and the median. For Hispanic students, the effects of TFA appear to exhibit a sharp decreasing trend across the math achievement distribution and the effect in fact becomes negative, although statistically insignificant, at the 90<sup>th</sup> quantile. Ideally, the positive effect of TFA would have shown itself across the entire math achievement distribution for

<sup>13</sup> For a detailed discussion of the gender gap in student achievement see Bedard and Cho (2010) and references therein.

Hispanic students however we do not find any evidence to suggest that TFA teachers, relative to control teachers, hurt Hispanic students in terms of their math test scores either.

In addition, we find evidence of heterogeneity in the effect of TFA for students taught by novice teachers; with the effects of TFA monotonically decreasing across the math achievement distribution. Allowing TFA teachers to teach in disadvantaged schools has the ability to have an extremely large impact, particularly at the bottom of the math achievement distribution given students in disadvantaged neighborhoods are generally taught by teachers with lower levels of experience (for evidence of this see Hanushek et al. 2004). Finally, we find that the positive effect of TFA is fairly homogeneous across the math achievement distribution irrespective of certification type.

Interestingly, a recent study by Constantine et al. (2009) compares the impact of traditionally certified teachers relative to less selective alternatively certified teachers (not including TFA teachers) on student achievement (math and reading) in disadvantaged schools (although somewhat less disadvantaged than schools in the TFA sample) using random assignment. They find that while less selective alternatively certified teachers do not hurt students they do not appear to help them either. Therefore, while RTTT calls for states to allow for alternative routes to certification, these findings illustrate that not all alternative routes appear to be created equal. As such, states may consider adopting programs that recruit teachers in a manner consistent with the criteria used in TFA in disadvantaged neighborhoods where turnover of teachers is high.

Taken together, these patterns provide tentative empirical support for the emphasis RTTT places on improving teacher effectiveness through alternative routes to certification and improving schools in disadvantaged neighborhoods. However in order for policies like RTTT to

be successful in improving student test score outcomes, understanding the full effects of alternative routes of certification, including but not limited to the TFA program, requires more empirical investigation. In particular, a number of questions remain unanswered. For instance, why does TFA differentially influence male and female student math achievement outcomes? What is the source of the heterogeneity for Hispanic students and for students taught by novice teachers? Is it possible to replicate the characteristics of the TFA program on a bigger scale? Addressing these questions would be a fruitful area for future research.

## References:

- Aaronson, Daniel, Lisa Barrow and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public Schools." *Journal of Labor Economics* 25(1): 95-135.
- Abadie, Alberto, Joshua D. Angrist, and Guido Imbens. 2002. "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings." *Econometrica* 70(1): 91-117.
- Abrevaya, J. and C. M. Dahl. 2008. "The Effects of Birth Inputs on Birthweight: Evidence From Quantile Estimation on Panel Data." *Journal of Business and Economic Statistics*, 26(4): 379-97.
- Antecol, Heather, Ozkan Eren, and Serkan Ozbeklik. 2013. "The Effect of Teacher Gender on Student Achievement in Primary School." IZA Discussion Paper 6453.
- Bedard, Kelly and Insook Cho. 2010. "Early Gender Test Score Gaps across OECD Countries." *Economics of Education Review*, 29, 348-363.
- Bitler, Marianne P., Jonah B. Gelbach, and Hillary W. Hoynes. 2006. "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments." *American Economic Review* 96(4): 988-1012.
- Bitler, Marianne P., Jonah B. Gelbach, and Hillary W. Hoynes. 2008. "Distributional Impact of Self Sufficiency Project." *Journal of Public Economics* 92(3-4): 748-65.
- Buchinsky, Moshe. 1994. "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression." *Econometrica*, 62(2): 405-58.
- Buchinsky, Moshe. 1998. "Recent Advances in Quantile Regression Models." *Journal of Human Resources*, 33(1): 88-126.
- Canay, Ivan A. 2011. "A Simple Approach to Quantile Regression for Panel Data." *The Econometrics Journal*, 14 (3): 368-86.
- Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources*, 41(4): 778-820.
- Clotfelter, Charles T., Elizabeth Glennie, Helen F. Ladd, and Jacob L. Vigdor. 2008. "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina." *Journal of Public Economics*, 92(5-6): 1352-1370.
- Constantine, Jill, Daniel Player, Tim Silva, Kristin Hallgren, Mary Grider, John Deke and Elizabeth Warner. 2009. "An Evaluation of Teachers Trained through Different Routes to

Certification: Final Report.” Institute of Education Sciences, NCEE 2009-4043, Washington, DC.

Darling-Hammond, Linda, Deborah J. Holtzman, Su Jin Gatlin, and Julian V. Heilig. 2005. “Does Teacher Preparation Matter? Evidence about Teacher Certification, Teach for America, and Teacher Effectiveness.” *Education Policy Analysis Archives*, 13(42).

Decker, Peter T., Daniel P. Mayer, and Steven Glazerman. 2004. “The Effects of Teach for America on Students: Findings from a National Evaluation.” Mathematica Policy Research Report 8792-8750, New York.

Eren, Ozkan and Serkan Ozbeklik. 2012. “Who Benefits from Job Corps? A Distributional Analysis of an Active Labor Market Program.” *Journal of Applied Econometrics*, forthcoming.

Fleischman, Howard L., Paul J. Hopstock, Marisa P. Pelczar, Brooke E. Shelley, and Brooke E. Shelley. 2010. *Highlights from PISA 2009: Performance of U.S. 15-Year-Old Students in Reading, Mathematics, and Science Literacy in an International Context* (NCES 2011-004).

Friedlander, Daniel and Philip K. Robins. 1997. “The Distributional Impacts of Social Programs.” *Evaluation Review* 21(5): 531-53.

Frölich, Markus and Blaise Melly. 2010. “Unconditional Quantile Treatment Effects under Endogeneity.” *Journal of Business and Economic Statistics*, forthcoming.

Goldhaber, Don D., and Dominic J. Brewer. 2000. “Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement.” *Educational Evaluation and Policy Analysis*, 22(2): 129-145.

Goldhaber, Don D., and Dominic J. Brewer. 2001. “Evaluating the Evidence on Teacher Certification: A Rejoinder.” *Educational Evaluation and Policy Analysis*, 23(1): 79-86.

Hanushek, Eric A., John F. Kain, and Steven Rivkin. 2004. “Why Public Schools Lose Teachers.” *Journal of Human Resources*, 39(2): 326-354.

Heckman, James J., Jeffrey A. Smith, and Nancy Clements. 1997. “Making the Most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts.” *Review of Economic Studies*, 64(4): 487-535.

Jackson, Erika and Marianne E. Page. 2013. “Estimating the Distributional Effects of Education Reforms: A Look at Project STAR.” *Economics of Education Review*, 32, 92-103.

Jepsen, Christopher and Steven G. Rivkin. 2009. “Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size.” *Journal of Human Resources*, 44(1): 223-250.

- Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2008. "What Does Certification Tell Us about Teacher Effectiveness? Evidence from New York City." *Economics of Education Review*, 27: 615-631.
- Koenker, R. 2004. "Quantile Regression for Longitudinal Data." *Journal of Multivariate Analysis*, 91: 74-89.
- Koenker, Roger W., and Gilbert Bassett. 1978. "Regression Quantiles." *Econometrica*, 46(1): 33-50.
- Koenker, Roger and Yannis Biliias. 2001. "Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments." *Empirical Economics* 26(1): 199-220.
- Krueger, Alan B. and Diane M. Whitmore. 2001. "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Results: Evidence from Project STAR." *Economic Journal*, 111(468): 1-28.
- Laczko-Kerr, Ildiko and David C. Berliner. 2002. "The Effectiveness of "Teach for America" and Other Under-Certified Teachers on Student Academic Achievement: A Case of Harmful Public Policy". *Education Policy Analysis Archives*, 10(37).
- Pilcher, Lorene C., and Donald C. Steele. 2005. "Teach for America and Regularly Certified Teachers: Teacher Efficacy, Teaching Concerns, Career Aspirations, and Teaching Effectiveness." In *Research on Alternative and Non-traditional Education*, p. 123-142 Lanham, Maryland: The Association of Teacher Educators.
- Powell, James L. 1986. "Censored Regression Quantiles." *Journal of Econometrics*, 32 (1): 143-55.
- Raymond, Margaret, Stephen H. Fletcher and Javier Luque. 2001. "Teach For America: An Evaluation of Teacher Differences and Student Outcomes in Houston, Texas." Stanford, CA: The Hoover Institution, Center for Research on Education Outcomes.
- Ripley, Amanda. 2010. "What Makes a Great Teacher." *The Atlantic*, January/February Issue.
- Rivkin, Steven G., Eric Hanushek, and John Kain. 2005. "Teachers, Schools and Academic Achievement." *Econometrica* Vol. 73, No. 2.
- Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review*, May Papers and Proceedings: 247-252.
- Rosen, Adam. 2012. "Set Identification via Quantile Restrictions in Short Panels." *Journal of Econometrics*, 166(1): 127-37.
- Scafidi, Benjamin, David L. Sjoquist and Todd R. Stinebrickner. 2007. "'Race, Poverty, and Teacher Mobility.'" *Economics of Education Review*, 26(2): 145-159.

**Table 1: Student Summary Statistics and Randomization Regressions**

	TFA Students		Control Students		Coefficient (Standard Error)
	Mean	S.D.	Mean	S.D.	TFA
<b>Math Test Score (NCE)</b>	32.82	19.20	31.61	19.87	.....
<b>Reading Test Score (NCE)</b>	26.36	20.26	26.87	20.75	.....
<b>Baseline Math Test Score (NCE)</b>	29.39	18.16	30.95	18.72	-0.001 (0.001)
<b>Baseline Reading Test Score (NCE)</b>	25.71	19.71	25.48	21.52	0.000 (0.001)
<b>Female</b>	0.50	0.50	0.48	0.49	0.015 (0.014)
<b>Race</b>					
White	0.02	0.16	0.02	0.16	-0.145 (0.122)
Black	0.72	0.44	0.63	0.48	0.155 (0.150)
Hispanic	0.23	0.42	0.29	0.45	-0.119 (0.108)
<b>Free/Reduced Lunch (%)</b>	97.82	14.15	98.14	13.49	-0.045 (0.109)
<b>Proportion of Students in Class that are not in Research Sample</b>	21.24	12.83	19.30	13.33	0.009 (0.008)
<b>Sample Size</b>	755		955		

NOTES: Randomization regression tests control for block fixed effects and the standard errors are clustered at the block level. TFA indicator takes the value of one if the student is taught by a TFA teacher and zero if the student is taught by a control teacher. See Appendix Table 1 for all other variable definitions.

**Table 2: Teacher Characteristics, Teaching Practices and Beliefs**

	Control Teachers		Novice Control Teachers		TFA Teachers	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
<b>Female</b>	0.86	0.34	0.83	0.38	0.68	0.47
<b>Teacher's Age</b>	38.85	11.57	31.72	8.22	25.80	4.09
<b>Teacher's Race</b>						
White	0.09	0.28	0.11	0.32	0.67	0.47
Black	0.75	0.43	0.55	0.51	0.20	0.4
Hispanic	0.13	0.34	0.27	0.46	0.05	0.22
<b>Traditional Teaching Certification</b>	0.63	0.48	0.33	0.48	0.44	0.50
<b>Teaching Practices</b>						
% of Time Spent on Academic Instruction in a Typical Day	73.82	13.06	74.22	10.61	72.19	15.78
% of Time Spent on Managing Classroom in a Typical Day	15.35	10.66	14.00	8.28	17.92	15.36
Students Working Independently in Small Groups (1=Yes)	0.98	0.13	0.94	0.23	0.95	0.21
Students Selecting Their Own Activities (1=Yes)	0.74	0.44	0.66	0.48	0.58	0.49
<b>Teacher Beliefs</b>						
% Who Place Major Emphasis on:						
Memorizing Facts, Rules and Steps	0.62	0.48	0.52	0.51	0.26	0.44
Getting the Right Answer	0.57	0.49	0.47	0.51	0.09	0.30
Performing Computation with Speed and Accuracy	0.22	0.42	0.16	0.38	0.39	0.49
<b>Sample Size</b>		58		18		41

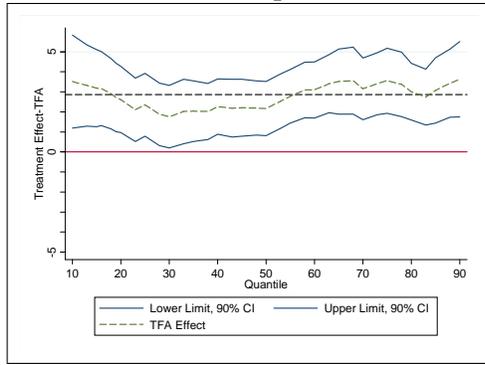
NOTES: See Appendix Table 1 for variable definitions.

**Table 3: Mean and Selected Fixed Effect Quantile Regression Estimates of TFA on Math Test Scores**

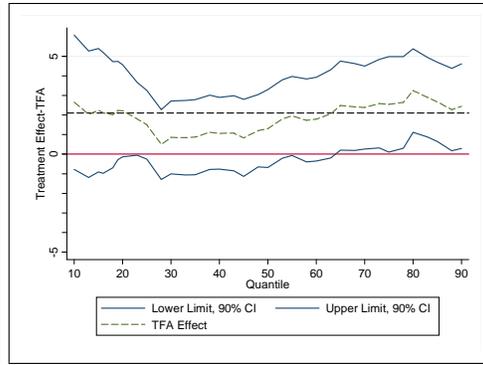
	Coefficient (Standard Error)					
	Mean	$\tau=10$	$\tau=25$	$\tau=50$	$\tau=75$	$\tau=90$
<b>Panel A: Full Sample</b>						
TFA Impact-Math Test Scores (Sample Size=1,710)	2.86** (0.82)	3.51** (1.40)	2.35** (0.95)	2.16** (0.81)	3.55* (0.98)	3.63* (1.14)
<b>Gender</b>						
<b>Panel B: Males</b>						
TFA Impact-Math Test Scores (Sample Size=870)	2.10* (1.20)	2.65 (2.08)	1.50 (1.06)	1.30 (1.20)	2.54* (1.48)	2.44* (1.31)
<b>Panel C: Females</b>						
TFA Impact-Math Test Scores (Sample Size=840)	3.69** (0.99)	3.47** (1.35)	3.48** (1.05)	4.21** (1.00)	4.45** (1.21)	2.92 (1.99)
<b>Race</b>						
<b>Panel D: Blacks</b>						
TFA Impact-Math Test Scores (Sample Size=1,149)	3.02** (1.01)	3.54** (1.95)	2.09* (1.2)	2.13* (1.16)	4.36** (1.21)	4.46** (1.15)
<b>Panel E: Hispanics</b>						
TFA Impact-Math Test Scores (Sample Size =460)	1.54 (1.37)	4.92** (2.40)	4.03** (1.84)	1.83 (1.29)	1.84 (1.44)	-2.25 (2.66)
<b>Teacher's Experience</b>						
<b>Panel F: Novice Teachers</b>						
TFA Impact-Math Test Scores (Sample Size=910)	2.89 (2.34)	6.88** (2.03)	3.39** (1.42)	2.52** (1.02)	2.06 (1.52)	0.14 (2.21)
<b>Traditional Teacher Certification</b>						
<b>Panel G: TFA vs. Certified Teachers</b>						
TFA Impact-Math Test Scores (Sample Size=1,370)	2.85** (1.00)	2.95** (1.46)	1.62 (1.11)	1.82** (0.90)	2.97** (1.06)	4.61** (1.08)
<b>Panel H: TFA vs. Uncertified Teachers</b>						
TFA Impact-Math Test Scores (Sample Size=1,129)	3.44** (1.69)	5.37** (1.88)	3.55** (1.20)	3.30** (0.94)	3.85** (1.50)	3.06* (1.83)

NOTES: All test scores are expressed in NCEs. The NCE scale has a mean of 50 and standard deviation of 21.06 nationally. Standard errors clustered at the block level are reported for mean effects. The standard errors for quantile regression estimates are obtained via 500 clustered (block level) bootstrap repetitions. Regression specifications in Panels A, F, G and H control for baseline test scores, student gender, student race/ethnicity, eligibility for free or reduced-price lunch and percentage of students in the classroom who were not in the research sample. Regression specification in Panels B and C exclude student gender from the set of control variables and the specifications in Panels D and E exclude student race/ethnicity. "\*\*\*" and "\*" indicate 5% and 10% significance levels, respectively.

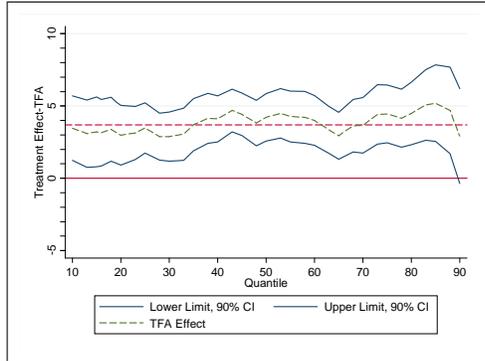
**Panel A: Full Sample**



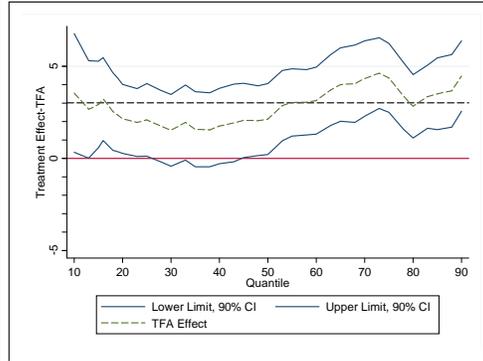
**Panel B: Males**



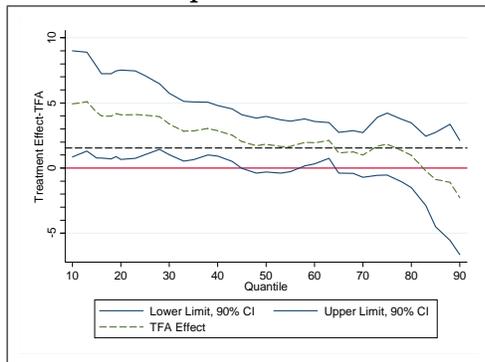
**Panel C: Females**



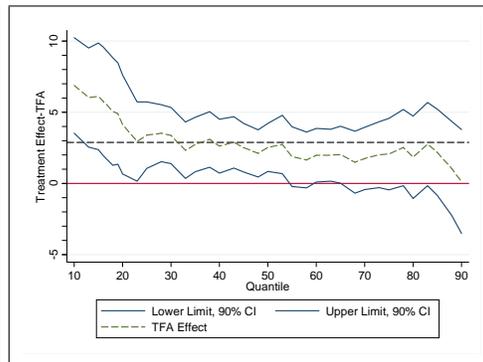
**Panel D: Blacks**



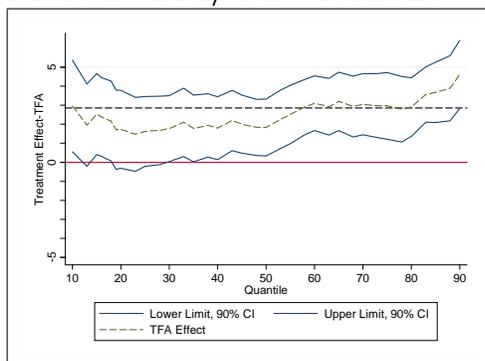
**Panel E: Hispanics**



**Panel F: Novice Teachers**



**Panel G: TFA/Cert. Teachers**



**Panel H: TFA/Uncert. Teachers**

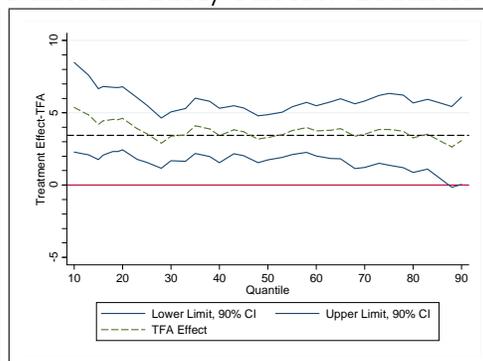


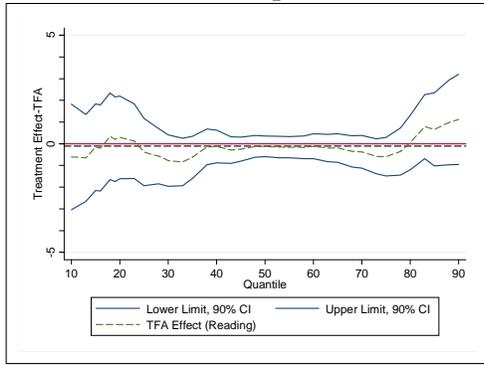
Figure 1: Fixed Effects Quantile Regression Estimates-Math Test Scores: The horizontal dashed lines are the mean effects.

The standard errors are obtained via bootstrapping based on 500 clustered (block level) repetitions.

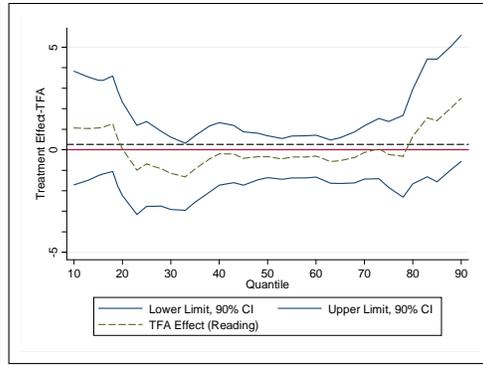
Appendix Table 1. Variable Definitions

Variable	Definition
<b>Student Characteristics:</b>	
Math Test Score (NCE)	Normal curve equivalent post-treatment math test score
Baseline Math Test Score (NCE)	Normal curve equivalent pre-treatment math test score
Female	=1 if student is female, and zero otherwise
White	=1 if student is non-Hispanic white and/or Asians/Pacific Islander, and zero otherwise
Black	=1 if student is non-Hispanic black, and zero otherwise
Hispanic	=1 if student is Hispanic, and zero otherwise
Free/Reduced Lunch	=1 if student had access to free/reduced lunch, and zero otherwise
Not in Research Sample	% of class that is not in research sample
<b>Teacher Characteristics:</b>	
TFA	=1 if TFA teacher, and zero if control teacher
Female	=1 if teacher is female, and zero otherwise
Age	Age of teacher
White	=1 if teacher is non-Hispanic white and/or Asians/Pacific Islander, and zero otherwise
Black	=1 if teacher is non-Hispanic black, and zero otherwise
Hispanic	=1 if teacher is Hispanic, and zero otherwise
Traditional Teacher Certification	=1 if teacher has traditional teacher certification, and zero if teacher has alternative teacher certification (i.e., emergency, temporary, initial, and other)
<i>Teaching Practices</i>	
Time Spent on Academic Instruction	% of Time Spent on Academic Instruction in a Typical Day
Time Spent on Managing Classroom	% of Time Spent on Managing Classroom in a Typical Day
Students Working Independently in Small Groups	=1 if students are working independently in small groups, and zero otherwise
Students Selecting Their Own Activities	=1 if students are selecting their own activities, and zero otherwise
<i>Teacher Beliefs</i>	
Memorizing Facts, Rules and Steps	=1 if teacher places a major emphasis on memorizing facts, rule and steps, and zero otherwise
Getting the Right Answer	=1 if teacher places a major emphasis on getting the right answer, and zero otherwise
Performing Computation with Speed and Accuracy	=1 if teacher places a major emphasis on performing computation with speed and accuracy, and zero otherwise

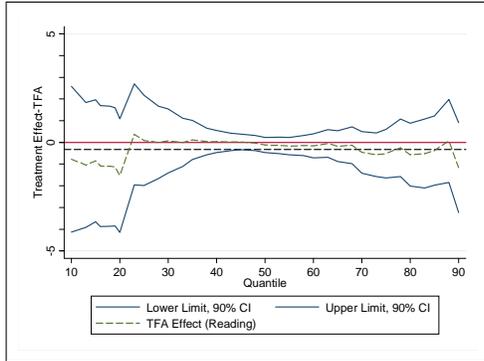
**Panel A: Full Sample**



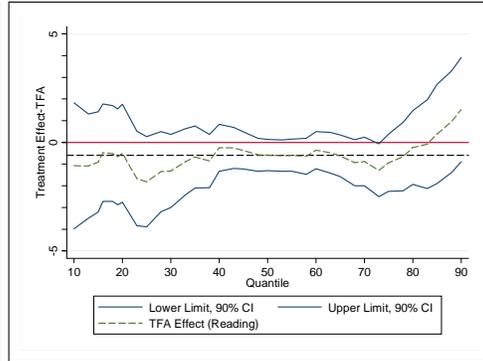
**Panel B: Males**



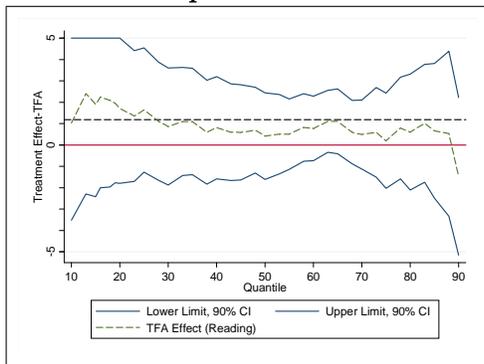
**Panel C: Females**



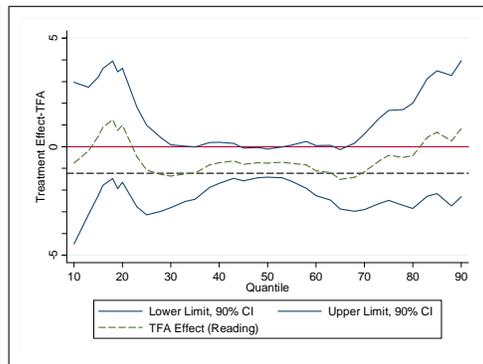
**Panel D: Blacks**



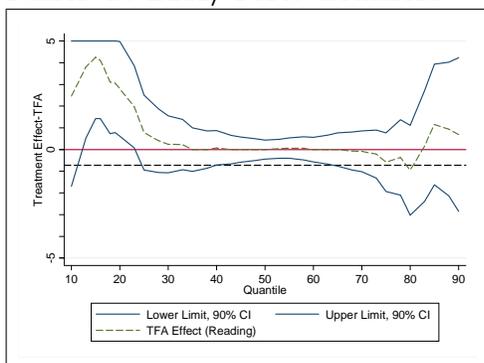
**Panel E: Hispanics**



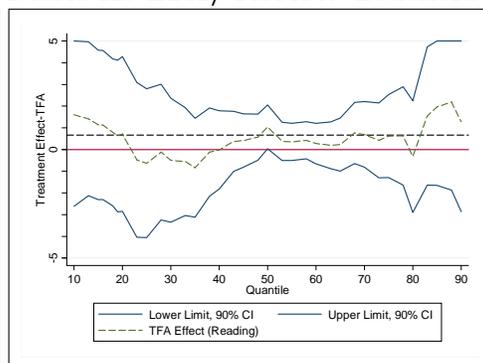
**Panel F: Novice Teachers**



**Panel G: TFA/Cert. Teachers**



**Panel H: TFA/Uncert. Teachers**



Appendix Figure 1: Fixed Effects Quantile Regression Estimates-Reading Test Scores: The horizontal dashed lines are the mean effects. The standard errors are obtained via bootstrapping based on 500 clustered (block level) repetitions.