The Impact of Voluntary Youth Service on Future Outcomes: Evidence from Teach For America*

Will Dobbie
Harvard University

Roland G. Fryer, Jr.
Harvard University and NBER

May 2013

Abstract

Nearly one million American youth have participated in service programs such as Peace Corps and Teach For America. This paper provides the first causal estimate of the impact of service programs on those who serve, using data from a web-based survey of former Teach For America applicants. We estimate the effect of voluntary youth service using a discontinuity in the Teach For America application process. Participating in Teach For America increases racial tolerance, makes individuals more optimistic about the life prospects of poor children, and makes them more likely to work in education. We argue that these facts are broadly consistent with the "Contact Hypothesis," which states that, under appropriate conditions, interpersonal contact can reduce prejudice.

*We are grateful to Cynthia Cho, Heather Harding, Brett Hembree, Wendy Kopp, Ted Quinn, Cynthia Skinner, and Andy Sokatch for their assistance in collecting the data necessary for this project. We also thank Lawrence Katz and seminar participants in the Harvard Labor Lunch for helpful comments and suggestions. Brad Allan, Vilisa Curto, Abhirup Das, Sara D’Alessandro, Ben Hur Gomez, Meghan Howard, Daniel Lee, Sue Lin, George Marshall, Rachel Neiger, Brendan Quinn, Wonhee Park, Gavin Samms, Jonathan Scherr, and Allison Sikora provided truly exceptional research assistance and project management support. Financial support from the Education Innovation Lab at Harvard University [Fryer], and the Multidisciplinary Program on Inequality and Social Policy [Dobbie] is gratefully acknowledged. Correspondence can be addressed to the authors by e-mail: dobbie@fas.harvard.edu [Dobbie] or rfryer@fas.harvard.edu [Fryer]. The usual caveat applies.
“We need your service, right now, at this moment in history.... I’m asking you to help change history’s course. Put your shoulder up against the wheel. And if you do, I promise you - your life will be richer, our country will be stronger, and someday, years from now, you may remember it as the moment when your own story and the American story converged, when they came together, and we met the challenges of our new century.”

President Barack Obama, at the signing of the Edward M. Kennedy Serve America Act

1 Introduction

Over the past half century, nearly one million American youth have participated in national service programs such as the Peace Corps, AmeriCorps, Teach For America, and City Year. These organizations have two stated objectives. The first is to provide services to communities in need. Peace Corps sends volunteers to work in education, business, information technology, agriculture, and the environment in more than 70 countries. Volunteers in Service to America (VISTA), an AmeriCorps program, enlists members to serve for a year at local nonprofit organizations or local government agencies. Teach For America (TFA) recruits recent, accomplished college graduates to teach in some of the most challenging urban public schools.

There is emerging empirical evidence that service organizations benefit the individuals that they serve. Decker et al. (2006) find that students randomly assigned to Teach For America corps members score 0.04 standard deviations high in reading and 0.15 standard deviations higher in math compared to students in classrooms with traditional teachers. Moss et al. (2001) find that students enrolled in an AmeriCorps tutoring program experience larger than expected gains in reading performance.

A second objective of service organizations is to influence the values and future careers of those who serve. Peace Corps stated mission includes helping to “promote a better understanding of other peoples on the part of Americans.” VISTA hopes to encourage its members to fight poverty throughout their lifetimes. Teach For America aims to develop a corps of alumni dedicated to ending educational inequity even after their two-year commitment is over. Advocates of service organizations point to notable alumni such as Christopher Dodd (Peace Corps), Reed Hastings (Peace Corps), and Michelle Rhee (Teach For America), as evidence of the long term impact on individuals who serve.

Despite nearly a million service program alumni and annual government support of hundreds of millions of dollars, there is no credible evidence of the causal impact of service on those who serve.2

---

1 This includes approximately 200,000 Peace Corps volunteers (http://multimedia.peacecorps.gov/multimedia/pdf/about/pc_facts.pdf), 637,000 AmeriCorps Volunteers (http://www.americorps.gov/about/newsroom/states_news.asp), 28,000 Teach For America corps members (http://www.teachforamerica.org/about-us/our-history/), and 13,700 City Year volunteers (http://www.cityyear.org/mediakit.aspx).

2 The 2010 federal budget included $373 million for Peace Corps and $98 million for AmeriCorps VISTA. Decker et al. (2006) report that school districts typically contribute about $1,500 per TFA member to offset recruiting costs,
This is due, in part, to the fact that service alumni likely had different values and career goals even before serving. As a result, simple comparisons of service program alumni and non-alumni are likely to be biased.

This paper provides the first causal estimate of the impact of service programs on those who serve, using data from a web-based survey of TFA applicants - both those that were accepted and those that were not - administered for the purposes of this study. The survey includes questions on an applicant’s background, educational beliefs, employment, political idealism, and racial tolerance. The section on educational beliefs asks about the extent to which individuals feel that the achievement gap is solvable, and the importance of teachers in reaching that goal. Employment variables measure whether individuals are interested in working in education in the future, whether they are currently employed in education, and whether they prefer to work in an urban or suburban school. Political idealism is captured through a series of questions such as whether or not the respondent self-identifies as liberal, or whether America should spend more money on specific social policies. Racial tolerance is captured using an Implicit Association Test. For a complete list of questions, see Online Appendix B.

Our identification strategy exploits the fact that admission into TFA is a discontinuous function of an applicant’s predicted effectiveness, calculated using a weighted average of scored responses to interview questions. As a result, there exists a cut-off point around which very similar applicants receive different application decisions. The crux of our identification strategy is to compare the average outcomes of individuals just above and below this cutoff. Intuitively, we attribute any discontinuous relation between average outcomes and the interview score at the cutoff to the causal impact of service in TFA.

The key threat to a causal interpretation of our estimates is that applicants may selectively respond to our survey. In particular, one may be concerned that TFA alumni will be more likely to respond, or that the non-alumni who respond will be different in some important way. Such selective response could invalidate our empirical design by creating discontinuous differences in respondent characteristics around the score cutoff. We evaluate this possibility in three ways. First, we test whether the survey response rate changes at the admissions cutoff. Second, we test whether the observable characteristics of survey respondents trend smoothly through the admissions cutoff score. Finally, we examine the density of survey respondents around the cutoff score. In all three cases, we find no evidence of the type of selective survey response that would invalidate our research design.

Our empirical analysis finds that serving in Teach For America increases an individual’s faith in education, increases an individual’s involvement in education, and increases racial tolerance. Polit-
ical idealism remains essentially unchanged. In particular, TFA alumni are 52.8 percentage points more likely to believe that the “achievement gap is a solvable problem,” and 48.0 percentage points more likely to believe that teachers are the most essential determinant of a student’s success. TFA alumni are also 35.5 percentage points more likely to work for a K-12 school, and 47.5 percentage points more likely to work in an education related career. Finally, serving in TFA increases implicit white-black tolerance by almost a full standard deviation. While service is also associated with higher implicit white-Hispanic tolerance, higher explicit white-black tolerance in the Modern Racism Scale, and a higher probability of believing that blacks and Hispanics are at least as intelligent as both whites and Asians, none of the estimates for explicit racial tolerance are statistically significant. Taken literally, this implies that while there is little treatment effect on self-reported measures of explicit tolerance, TFA increases the unconscious tolerance of its members.

Analysis of subsamples reveals that the impact of service on faith in education and educational involvement is larger for men, with no systematic differences by ethnicity or Pell Grant receipt – a rough proxy for household income at college entry.

We argue that this new set of facts, particularly those on racial beliefs, is consistent with the “Contact Hypothesis,” developed by Allport (1954) and extended by Brown and Hewstone (2005), Hewstone and Brown (1986), and Pettigrew (1998). The “Contact Hypothesis” states that intergroup contact increases tolerance. The majority of the empirical evidence shows that intergroup contact is negatively correlated with intergroup prejudice (Pettigrew and Tropp 2006), and in real world experimental settings, Van Laar et al. (2005) and Boisjoly et al. (2006) show that living with a minority group increases tolerance among white college students, while, and Barnhardt (2009) shows that living with a minority group increases the tolerance of Hindu children. Similarly, Clingingsmith et. al (2009) demonstrate that winning a lottery to participate in the Hajj pilgrimage to Mecca increases belief in equality and harmony of ethnic groups among other outcomes.

TFA service typically involves sending a college educated young adult, whose parental income is above the national average, into a predominantly poor and minority neighborhood. Seventy-four percent of corps members in our survey sample are white, and eighty percent have at least one parent with a college degree. The average parental income of a corps member while in high school is $118 thousand, compared to the national median family income of approximately $50 thousand (U.S. Census Bureau 2000). In sharp contrast to this privileged upbringing, roughly eighty percent of the students taught by corps members qualify for free or reduced-price lunch and more than ninety percent are African-American or Hispanic. To the best of our knowledge, our analysis is the first to estimate the impact of contact with a poor and minority group on the beliefs of more advantaged individuals.

There are three potentially important caveats to our analysis. First, because TFA introduced its discontinuous method of selecting applicants in 2007, our primary analysis includes only one cohort of TFA applicants surveyed roughly a year after their service commitment ended. To address this issue, we also collected data on TFA applicants from the 2003 to 2006 cohorts. Applicants in these cohorts were admitted only if they met prespecified interview subscore requirements. For
example, TFA admitted applicants with the highest possible interview score in perseverance and organizational ability so long as they had minimally acceptable scores in all other areas. In total, there were six separate combinations of interview subscores that met the admissions requirements. We estimate the impact of service for the 2003 to 2006 cohorts by instrumenting for TFA placement using an indicator variable equal to one if a candidate meets one of the six subscore criteria for admissions, controlling for fully nonparametric controls for each individual interview subscore. The impact of TFA service is therefore identified using the interaction of the subscores. Our key identifying assumption is that, conditional on our nonparametric controls, the that the interaction of interview subscores only impacts future outcomes through TFA placement. These estimates suggest that the impacts of service are persistent, with older TFA alumni being more likely to believe in the power of education, more likely to be involved or employed in education, and more racially tolerant.

A second caveat is that the response rate of the 2007 cohort to our web-based survey is only 30 percent. While there is no evidence of the type of selective survey response around the interview score cutoff that would invalidate our empirical design, we cannot rule out unobserved differences in who responded to our survey. With such a low response rate, the bounds on our analysis include both positive and negative effects, which is typical in web-based surveys. For instance, approximately 26 percent of University of Chicago Business School alumni from the graduating classes of 1990 to 2006 responded to a web-based survey conducted in Bertrand, Goldin and Katz (2010); 10 percent of individuals receiving UI benefits in New Jersey initially responded to an online survey conducted in Krueger and Mueller (2011), with an additional 60 percent of those initial respondents dropping out of the follow-up survey.

Third, although TFA is broadly similar to other service organizations, it also differs in important ways that limit our ability to generalize our results. To the extent that TFA’s impact on alumni is driven by factors that all service organizations have in common, the results of our study will be informative about the effects of service programs more generally. If one believes that the unique attributes of TFA such as its selectivity or focus on urban teaching drive its impact, the results of the study should be interpreted more narrowly.

The paper proceeds as follows. Section ?? provides a brief overview of Teach For America and its relationship to other prominent service programs around the world. Section ?? describes our web-based TFA survey and sample. Section ?? details our research design and econometric framework for estimating the causal impact of TFA on racial and educational beliefs, employment outcomes, and political idealism. Section ?? describes our results. The final section concludes. There are three online appendices. Online Appendix A provides further details of how we coded variables used in our analysis and constructed the samples. Online Appendix B provides implementation details and the complete survey administered to TFA applicants. Online Appendix C provides additional results and robustness checks of our main analysis excluded from the main text.
2 A Brief Overview of Teach For America

A. History
Teach For America (TFA), a non-profit organization that recruits recent college graduates to teach for two years in low-income communities, is one of the nation’s most prominent service programs. Based on founder Wendy Kopp’s undergraduate thesis at Princeton University, TFA was created to build a movement to eliminate educational inequity by enlisting our nation’s most promising future leaders as teachers. In 1990, TFA’s first year in operation, Kopp raised $2.5 million and attracted 2,500 applicants for 500 teaching slots in New York, North Carolina, Louisiana, Georgia, and Los Angeles.

Since its founding, TFA corps members have taught more than three million students. Today, there are 8,200 TFA corps members in 125 “high-need” districts across the country, including 13 of the 20 districts with the lowest graduation rates. Roughly 80 percent of the students reached by TFA qualify for free or reduced-price lunch and more than 90 percent are African-American or Hispanic.

B. Application Process
Entry into TFA is highly competitive; in 2010, more than 46,000 individuals applied for just over 4,000 spots. Twelve percent of all Ivy League seniors applied. A significant number of seniors from historically black colleges and universities also applied, including 1 in 5 at Spelman College and 1 in 10 at Morehouse College. TFA reports that 28 percent of incoming corps members received Pell Grants, and almost one-third are people of color.

In its recruitment efforts, TFA focuses on individuals who possess strong academic records and leadership capabilities, regardless of whether or not they have exposure to teaching practice prior to entry into TFA. Despite this lack of formal training, students assigned to TFA corps members score about 0.15 standard deviations higher in math and 0.04 standard deviations higher in reading than students assigned to traditionally certified teachers (Decker et al. 2006).

To apply, candidates complete an online application, which includes a letter of intent, and a resume. After a phone interview, the most promising applicants are invited to participate in an in-person interview, which includes a sample teaching lesson, a group discussion, a written exercise, and a personal interview. Applicants who are invited to interview are also required to provide transcripts, obtain two on-line recommendations, and provide one additional reference.

Using information collected through the application and interview, TFA bases their selection of candidates on a model that accounts for multiple criteria that they believe are linked to success in the classroom, including achievement, perseverance, critical thinking, organizational ability, motivational ability, respect for others, and commitment to the TFA mission. TFA conducts ongoing research on their selection criteria, focusing on the link between the selection criteria and observed single-year gains in student achievement in TFA classrooms.

Between 2003 and 2006, TFA admitted candidates who met prespecified interview subscore requirements. For example, TFA admitted all applicants with the highest possible achievement
score so long as they had minimally acceptable scores in all other areas. Or, applicants could be admitted by having the highest possible score in perseverance and organizational ability, again so long as they had minimally acceptable scores in all other areas. In total, there were six separate combinations of interview subscores that met the admissions requirements. In 2007, TFA conducted a systematic review of their admissions measures, improving the correlation between these subscores and internal TFA measures of classroom success. Also starting in 2007, TFA began using a linear function of the application subscores to help rank and select candidates.

C. Training and Placement

TFA cohorts included in our study were required to take part in a five-week TFA summer institute to prepare them for placement in the classroom at the end of the summer. The TFA summer institute includes courses covering teaching practice, classroom management, diversity, learning theory, literacy development, and leadership. During the institute, groups of participants also take full teaching responsibility for a class of summer school students.

At the time of their interview, applicants submit their subject, grade, and location preferences. TFA works to balance these preferences with the needs and requirements of districts. With respect to location, applicants rank each TFA region as highly preferred, preferred, or less preferred and indicate any special considerations, such as the need to coordinate with a spouse. Over 90 percent of the TFA applicants accepted are matched to one of their “highly preferred” regions (Decker et al., 2006).

TFA also attempts to match applicants to preferred grade levels and subjects, depending on an applicants’ academic backgrounds, district needs, and state and district certification requirements. As requirements vary from region to region, applicants may not be qualified to teach the same subjects and grade levels in all regions. It is also difficult for school regions to predict the exact openings they will have in the fall, and late changes in subject or grade-level assignments are not uncommon.

TFA corps members are employed and paid directly by the school districts for which they work, and generally receive the same salaries and health benefits as other first year teachers. Most districts pay a $1,500 per corps member fee to TFA to offset screening and recruiting costs. TFA gives corps members various additional financial benefits, including “education awards” of $4,725 for each year of service, which they can use toward past or future educational expenses, and transitional grants and no-interest loans to help corps members make it to their first paycheck.

TFA corps members are hired to teach in local school districts through alternative routes to certification. Typically, they must take and pass exams required by their districts before they begin teaching. Corps members may also be required to take additional courses to meet state certification requirements or to comply with the requirements for highly qualified teachers under the No Child Left Behind Act.
3 Teach For America Survey and Sample

To understand the impact of TFA on racial and educational beliefs, employment outcomes, and political idealism, we conducted a web-based survey of the 2003-2009 TFA application cohorts between April 2010 and May 2011. The survey contained 87 questions and lasted approximately 30 minutes. As an incentive to complete the survey, every individual was entered into a lottery for a chance to win $5,000. The complete survey is available in Online Appendix B.

A. Contacting TFA Applicants

Applicants were first contacted using the email addresses they supplied to TFA in their initial applications. Between April 2010 and June 2010, applicants received up to three emails providing them with information about the survey and a link to the survey. Each email reminded applicants that by completing the survey they would be automatically entered in a lottery for $5,000. Approximately 30 percent of the 2,567 2007 TFA alumni and 14 percent of the 4,801 2007 non-alumni started the survey during this phase. To increase the response rate among non-alumni, we also contacted individuals using phone numbers from TFA application records. We began by calling each of the non-alumni who had not responded to the email and who had provided TFA a valid phone number using an automated call system with a brief 30 second recording with information about the survey. We then contacted non-respondents using personal calls from an outsourced calling service. Voicemails were left for those who did not answer the phone. For the 2007 cohort, those who did not answer the phone were called again a few weeks later. This strategy resulted in an additional 589 responses among 2007 non-alumni. The process was similar for the 2003 - 2006 and 2008 - 2009 cohorts, though we made fewer follow up calls than with the 2007 cohort as these cohorts were not a priority for the analysis. Appendix B provides additional details on each step of this process.

These strategies yielded a final response rate of 32.7 percent among 2007 TFA alumni and 29.6 percent among 2007 non-alumni. Among the other cohorts, the response rate is lower for older cohorts and non-alumni. The difference in the response rate between alumni and non-alumni is smallest in the 2007 cohort, likely due to the additional phone calls to the non-alumni in this cohort. Response rates are presented for all cohorts in Appendix Figure 1.

One important threat to our identification strategy is that the response rate changes discontinuously at the score cutoff. In this scenario, our results may be driven by changes in the type of respondent at the cutoff rather than the causal impact of TFA. We test for differences in the response rate at the admissions cutoff in Figure 1. We plot the fraction of applicants taking the survey and fitted values from a regression of an indicator variable for taking the survey on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with scoring above the cutoff. There is not a statistically significant impact of scoring above the cutoff score on survey response. If anything, candidates with scores above the cutoff are somewhat less likely to have taken the survey. We test for differences in survey response at the admissions cutoff in Figure 1.

---

4Between June 2010 and July 2010 we also searched for both alumni and non-alumni non-responders on Facebook using accounts specifically created for the TFA survey. Account pages included an overview of and link to the survey. Despite finding 2,610 applicants on Facebook, this method yielded only 53 accepted “friend” requests.
response around the cutoff more formally in Section 5.4, finding no evidence that our identifying assumption is violated.\textsuperscript{5}

Another important concern is that the non-alumni members who do respond will be different from the alumni who respond in some important way. This concern is particularly relevant given that we recruited non-alumni members using both email and phone call strategies, while we recruited alumni using email strategies only. If responding non-alumni members are different than responding alumni, there will be discontinuous differences in respondent characteristics around the score cutoff that will invalidate our empirical design. In Section 5.4, we test this possibility by examining respondent characteristics around the score cutoff, finding no evidence of selective survey response. We also reestimate our main results using non-alumni members who responded to the email strategies only, finding nearly identical results to those with the full sample of respondents.

\textit{B. The Survey}

Data collected in our online survey of TFA applicants is the heart of our analysis. We asked applications about their demographics and background information, educational beliefs, employment outcomes and aspirations, political idealism, and racial beliefs. Whenever possible, survey questions were drawn from known instruments such as the College and Beyond Survey, Harvard and Beyond Survey, the The National Longitudinal Study of Adolescent Health Teacher Survey, the Modern Racism Scale, and the General Social Survey. In this paper, we use only a small fraction of the data we collected. For further details on these variables or those omitted from our analysis, see online Appendix B.

The set of questions on educational beliefs were designed to measure the extent to which individuals feel that the achievement gap is solvable, that schools can achieve that goal, and the importance of teachers in increasing student achievement. Survey respondents were asked whether they agreed or disagreed with a series of statements on a five point Likert scale ranging from “agree strongly” to “disagree strongly.” The questions used are similar to those asked in the The National Longitudinal Study of Adolescent Health Teacher Survey. Other, more open-ended, questions included “what fraction of blacks can we reasonably expect to obtain a college degree,” and “who is the most important in determining how well students perform in school?” For questions with answers that do not have clear cardinality, we create indicator variables equal to one if the response was “favorable” (e.g. strongly agree that the achievement gap is a solvable problem).

Employment variables measure whether individuals are interested in working in education in

\textsuperscript{5}Appendix Figure 2 presents point estimates and associated 95 percent confidence intervals for the reduced form impact of having a score above the cutoff (2007 - 2009) or having a set of eligible interview subscores (2003 - 2006) on survey response in all cohorts. For the 2007 - 2009 cohorts we control for interview score and interview score interacted with scoring above the cutoff score. For the 2003 - 2006 cohorts we include fully nonparametric controls for each interview subscore, and report the point estimate for an indicator variable equal to one if the combination of interview subscores is eligible for TFA admission. While there is little difference in the response rate of eligible and ineligible applicants from the 2007 - 2009 cohorts, eligible applicants from the 2003 - 2006 are 6 to 11 percentage points more likely to respond to our survey. In Section 5.4 we test whether the eligible applicants who take the survey are different on observable characteristics, finding that eligible applicants who took the survey are more likely to be white or Asian in the 2003 - 2006 cohorts. The results for the 2003 - 2006 cohorts should be interpreted with these differences in response rate in mind.
the future, whether they are currently employed in education, and whether they prefer to work in an urban or suburban school. Political idealism is captured in a series of questions such as whether or not the respondent self identifies as liberal or whether the government should spend more or less on issues such as closing the achievement gap, welfare assistance, and fighting crime. For political idealism, we create indicator variables equal to one if the response is more liberal.

In the final portion of the survey, we asked participants to take a 10 minute Implicit Association Test (IAT) that measured white-black implicit bias. We measure implicit associations as they better indicate unconscious feelings about minorities (Bertrand, Chugh, Mullainathan 2005). Manipulation has been shown to be harder on the IAT than other measures (Steffens 2004), and a recent meta-analysis reports the IAT is better at predicting behaviors than explicit attitudes when the association measured involves a white-black comparison (Greenwald et al. 2009). IAT scores also correlate well with other implicit measures of racial attitudes and real-world actions. For instance, individuals with more anti-black IAT scores are more likely to make negative judgments about ambiguous actions by blacks (Rudman and Lee 2002); more likely to exhibit a variety of micro-behaviors indicating discomfort with minorities, including less speaking time, less smiling, fewer extemporaneous social comments, more speech errors, and more speech hesitations in an interaction with a black experimenter (McConnell and Leibold 2001); and are more likely to show greater activation of the amygdala, an area of the brain associated with fear-driven responses, to the presentation of unfamiliar black versus white faces (Phelps et al. 2000). Moreover, in Sweden IAT scores predict discrimination in the hiring process among managers (Rooth 2007) and in the U.S., doctors with stronger anti-black IAT scores are less likely to prescribe thrombolysis for myocardial infarction to otherwise similar black patients (Green et al. 2006), though that latter finding has been questioned (Dawson and Arkes 2008).

We use a brief format IAT, developed by Sriram and Greenwald (2009), to assess the relative strength of automatic associations between “good” and “bad” outcomes and white and black faces. The brief format IAT performs similarly on test-retest and implicit-explicit correlations as the standard IAT, but with one third the number of trials. We standardize Implicit Association Test scores to have a mean of zero and a standard deviation of one, with higher values indicating less racial bias.

To complement the IAT measure of implicit bias, individuals were also asked about explicit racial bias. Our first measure of explicit bias comes from the General Social Survey. Individuals were asked to separately rate the intelligence of Asians, blacks, Hispanics, and whites on a seven point scale that ranged from “almost all are unintelligent” to “almost all are intelligent.” We recoded this variable to indicate whether individuals believe that blacks and Hispanics are at least as intelligent as whites and Asians. Our second measure of explicit bias is the Modern Racism Scale (McConahay 1983). The Modern Racism Scale consists of six questions with which individuals are asked how much they agree or disagree. Each item was re-scaled so that lower numbers are associated with a more

---

6Some critics argue that the IAT may be assessing shared norms, familiarity, perceptual salience asymmetries, or cultural knowledge that does not correspond to personal endorsement of that knowledge (e.g. Karpinski and Hilton 2001; Rothermund and Wentura 2004).
anti-black response, then a simple average was taken of the six questions. We normalized this scale to have mean zero and standard deviation one across each cohort. The six statements that individuals are presented are: “over the past few years, blacks have gotten more economically than they deserve;” “over the past few years, the government and news media have shown more respect for blacks than they deserve;” “it is easy to understand the anger of black people in America;” “discrimination against blacks is no longer a problem in the United States;” “blacks are getting too demanding in their push for equal rights;” and “blacks should not push themselves where they are not wanted.”

Index variables for each survey domain were also constructed by standardizing the sum of individual questions to have a mean of zero and a standard deviation of one in each cohort. Rather than add dichotomous and standardized variables together, we converted all standardized variables to indicator variables equal to one if the continuous version of the variable was above the median of the full sample. Results are qualitatively similar if we combine the original dichotomous and continuous variables instead. Details on the coding of each measure are available in Appendix A.

C. The Final Sample

Our final sample consists of data from our web-based survey merged to administrative data from Teach For America. The administrative records consist of admissions files and placement information for the 2003 to 2009 application cohorts. Our data includes all TFA applicants who attended the in-person interview. A typical applicant’s data include her name, undergraduate institution, GPA, and major, admissions decision, placement information, and interview score. Data from TFA administrative records and our web-based survey were matched using name, application year, college and email address. Our primary sample consists of all 2007 applicants who responded to our survey. Our secondary sample consists of survey respondents from all cohorts.

Summary statistics for the 2007 survey cohort are displayed in Table 1. Seventy-three percent of the sample is white, 7 percent is Asian, 7 percent is black, and another 6 percent is Hispanic.7 Alumni had an average college GPA of 3.58 while non-alumni had an average GPA of 3.46. The parents of both the typical alumni and non-alumni are highly educated. Fifty-eight percent of alumni have at least one parent with a graduate degree, while 50 percent of non-alumni do. Over 80 percent of both groups have at least one parent with a college degree. With that said, a significant fraction of TFA applicants come from disadvantaged backgrounds. Nineteen percent of TFA alumni in our sample were eligible for a Pell Grant in college, while 23 percent of non-alumni were eligible.

4 Research Design

Our identification strategy exploits the fact that entry into TFA is a discontinuous function of an applicant’s interview score. Consider the following model of the relationship between future

7The racial distribution of TFA applicants mirrors that of colleges graduates from selective colleges more broadly. Five percent of graduating seniors at “more selective” or “most selective” colleges are black. Six percent are Hispanic (U.S. Department of Education 2010). The 2011 TFA cohort is more diverse than previous cohorts, with 12 percent of the cohort identifying as black and 8 percent identifying as Hispanic.
outcomes \((y_i)\) and serving in TFA \((TFA_i)\):

\[
y_i = \alpha + \gamma TFA_i + \varepsilon_i
\]  

(1)

The parameter of interest is \(\gamma\), which measures the causal effect of service on future outcomes \(y_i\). The problem for inference is that if individuals select into service organizations because of important unobserved determinants of later outcomes, such estimates may be biased.\(^8\) In particular, it is plausible that people who select into service organizations had different beliefs and outcomes before they served: \(E[\varepsilon_i|y_{it-1}] \neq 0\). Since \(TFA_i\) may be a function of past beliefs and outcomes, this can lead to a bias in the direct estimation of \(\gamma\) using OLS. The key intuition of our approach is that this bias can be overcome if the distribution of unobserved characteristics of individuals who were just below the bar for TFA is the same as the distribution among those who were just above the bar:

\[
E[\varepsilon_i|\text{score}_i = c^* + \Delta]_{\Delta \to 0^+} = E[\varepsilon_i|\text{score}_i = c^* - \Delta]_{\Delta \to 0^+}
\]  

(2)

where \(\text{score}_i\) is an individual’s interview score, and \(c^*\) is the cutoff score below which very few applicants are admitted to TFA. Equation (2) implies that the distribution of individuals to either side of the cutoff is as good as random with respect to unobserved determinants of future outcomes \(\varepsilon_i\). In this scenario, we can control for selection into TFA using an indicator variable for whether an individual has an interview score above the cutoff as an instrumental variable. Since service in \(TFA_i\) is a discontinuous function of interview score, whereas the distribution of unobservable determinants of future outcomes \(\varepsilon_i\) is by assumption continuous at the cutoff, the coefficient \(\gamma\) is identified. Intuitively, any discontinuous relation between future outcomes and the interview score at the cutoff can be attributed to the causal impact of service in TFA under the identification assumption in equation (2).

Formally, let TFA placement \((TFA_i)\) be a smooth function of an individual’s interview score \((\text{score}_i)\) with a discontinuous jump at the eligibility cutoff \(c^*\):

\[
TFA_i = f(\text{score}_i) + \eta 1(\text{score}_i \geq c^*) + \varepsilon_i
\]  

(3)

In practice, the functional form of \(f(\text{score}_i)\) is unknown. We follow Card et al. (2008) and Lee and Lemieux (2010), among others, and approximate \(f(\text{score}_i)\) with a quadratic in interview score that is allowed to vary on either side of the cutoff. Estimation with a local linear regression, such as that presented in Ludwig and Miller (2007) and Almond et al. (2010), gives similar results (see Appendix Table 1). To address potential concerns about discreteness in the interview score in both our first and second stage results, we cluster our standard errors at the interview score level (Card and Lee 2008).

The identified second stage parameter measures the average treatment effect for individuals induced into TFA by scoring just above the cutoff. The key threat to a causal interpretation of our

---

\(^8\)Previous studies that examine the association between service and future outcomes, such as Jastrzab and Winship (2008) and McAdam and Brandt (2009), estimate equations such as (1).
estimates is that applicants may selectively respond to our survey:

\[ E[\varepsilon_i | \text{survey}_i = 1, \text{score}_i = c^* + \Delta]_{\Delta \to 0^+} \neq E[\varepsilon_i | \text{survey}_i = 1, \text{score}_i = c^* - \Delta]_{\Delta \to 0^+} \]

where \( \text{survey}_i = 1 \) represents the set of TFA applicants who answered at least one question in our web-based survey.

In particular, one may be concerned that former TFA alumni will be more likely to respond than non-alumni, or that the non-alumni who respond will be different in some important way from the alumni that respond. Such selective response could invalidate our empirical design by creating discontinuous differences in respondent characteristics around the score cutoff. In Section 5.4 we evaluate this possibility in three ways: by testing whether the observable characteristics of survey respondents trend smoothly through the cutoff, by examining whether the fraction of applicants taking the survey trends smoothly through the cutoff, and by examining the density of survey respondents around the cutoff. None of these tests points to evidence of the type of selective survey response that would invalidate the empirical design.

A separate concern is that there are discontinuous differences in the TFA interview subscores around the score cutoff. In this scenario, our results could be driven by preexisting differences in TFA candidates. We evaluate this possibility in Appendix Figure 3 by plotting each interview subscore against the total interview score for the entire 2007 applicant pool. Candidates with interview scores above the cutoff have higher critical thinking subscores, but there are no significant differences in achievement, perseverance, organizational ability, motivational ability, respect for others, and commitment to the TFA mission subscores. Results are identical if we restrict to 2007 TFA applicants who took the survey. These results suggest that, consistent with our empirical design, neither the TFA interview process or any kind of selective survey response based on interview score drives the results.

One final problem unique to our setting is that the cutoff score \( c^* \) must be estimated from the data. TFA does not specify a cutoff score each year. Rather, they select candidates using the interview score as a guide until a prespecified number of teaching slots are filled. Our goal is to identify the unknown score cutoff that best fits the data. To identify this optimal discontinuity point, we use a technique similar to that used to identify structural breaks in time series data and identify discontinuities in the dynamics of neighborhood racial composition (Card, Mas, and Rothstein 2008). Specifically, we regress an indicator variable equal to one if the individual was selected for TFA on a constant and an indicator variable equal to one if the individual scored above a particular cutoff \( c \) in the full sample of applicants. We then loop over all possible cutoffs \( c \) in 0.0001 intervals, selecting the value of \( c \) that maximizes the \( R^2 \) of our specification. Hansen (2000) shows that this procedure yields a consistent estimate of the true discontinuity. A standard result in the structural break literature (e.g., Bai 1997) is that one can ignore the sampling error in the location of the discontinuity when estimating the magnitude of the discontinuity. Using different cutoff points around the optimal \( c^* \) yield very similar results.
5 Results

5.1 First Stage

First-stage results of the impact of the score cutoff on TFA service are presented graphically in Figure 2. The figure presents actual and fitted values for our first stage regression. The sample includes all 2007 applicants to TFA who answered at least one question on our survey. An individual is defined as having served in TFA if she accepted the TFA offer and was assigned a school district. Actual values are plotted in bins of size 0.0025. The fitted values are from a regression of an indicator variable equal to one if an individual served in TFA on an indicator variable for being above the cut-off score, \( c^* \), a quadratic trend in interview score, and a quadratic trend in interview score interacted with the indicator for being above the cut-off score:

\[
TFA_i = \alpha_1 + \alpha_2 I(score_i \geq c^*) + \alpha_3 score_i + \alpha_4 score_i^2 + \alpha_5 score_i \cdot 1(score_i \geq c^*) + \alpha_6 score_i^2 \cdot 1(score_i \geq c^*) + \epsilon_i \tag{4}
\]

where \( \alpha_2 \) is the effect of having an interview score above the cutoff score on the probability of service.

Figure 2 suggests that there is a large and precisely estimated increase in the probability of serving in TFA at the cutoff score. About 35 percent of applicants scoring just below the cutoff serve, while approximately 65 percent of applicants scoring just above the threshold serve. The corresponding estimates are significant at the 1 percent level, suggesting that our empirical design has considerable statistical power.

5.2 Main Results

Figure 3 summarizes our main results, and Figures 4 through 7 present results for each set of questions separately. Each figure presents actual and fitted values for 2007 applicants to TFA who answered at least one question on our survey. Section ?? provides results for additional cohorts. Actual values are plotted in bins of size 0.0025. The fitted values are from a regression of the dependent variable on an indicator variable for being above the cut-off score, \( c^* \), a quadratic in interview score, and a quadratic in interview score interacted with the indicator for being above the cut-off score. Thus, similar to our first stage specification, our reduced form specification is:

\[
y_i = \alpha_1 + \alpha_2 I(score_i \geq c^*) + \alpha_3 score_i + \alpha_4 score_i^2 + \alpha_5 score_i \cdot 1(score_i \geq c^*) + \alpha_6 score_i^2 \cdot 1(score_i \geq c^*) + \epsilon_i \tag{5}
\]

Lee and Lemieux (2010) propose a formal test for optimal bin width based on the idea that if the bins are narrow enough, then there should not be a systematic relationship between the outcome variable and the running variable within each bin. Otherwise, the bin is too wide and the mean value of the outcome variable is not representative at the boundaries. A simple test for this consists of adding a set of interactions between the bin dummies and the running variable to a base regression of the outcome variable on the set of bin dummies, and testing whether the interactions are jointly significant. A bin width of 0.0025 passes this test.
where $\alpha_2$ is the reduced form effect of having an interview score above the cutoff score on the probability of service. Below each figure we present the p-value associated with the reduced form effect $\alpha_2$. Results in tabular form, including two stage least squares estimates, are available in Appendix C.

Our summary results in Figure 3 examine the standardized sum of dichotomous individual questions from that domain. Continuous variables were converted to dichotomous variables by creating an indicator variable equal to one if the continuous version of the variable was above the median of the full sample. As previously mentioned, results are qualitatively similar if we combine the dichotomous and continuous variables instead. The racial tolerance index is made up of an individual’s white-black IAT score only. Details on the coding of each measure are available in Appendix A. Figure 3 suggests that serving in Teach For America increases an individual’s faith in education, an individual’s involvement in education, and an individual’s racial tolerance. Individuals who serve score 1.582 standard deviations higher on our index of educational faith, and 1.880 standard deviations higher on our index of educational employment. TFA alumni are also 0.980 standard deviations “more tolerant” than non-alumni. Political idealism remains essentially unchanged.

Figure 4 presents results for each education belief variable separately. Serving in TFA increases an individual’s faith in the ability of poor children to compete with more advantaged children, and belief in the importance of teachers in raising student achievement. Two stage least squares estimates suggest that individuals who serve are 52.8 percentage points more likely to believe that poor children can compete with more advantaged children, 47.7 percentage points more likely to believe that the achievement gap is solvable, 48.0 percentage points more likely to believe that teachers are the most important determinant of success, 25.6 percentage points more likely to believe that schools can close the achievement gap without the help of families, and 84.5 percentage points more likely to disagree that there is little teachers can do to ensure that students succeed. Individuals who serve also believe that we can reasonably expect 30.3 percent more minorities to graduate from a four year college than individuals who do not serve.

The effect of TFA on involvement in education is depicted in Figure 5. An important criticism of TFA is that corps members frequently depart before their two-year commitment has been fulfilled or immediately after (Darling-Hammond et al. 2005). Our results do not address the question of whether TFA teachers are more likely to stay in education compared to other teachers. Instead, we ask whether TFA leads individuals to stay in education longer than they otherwise would have without TFA.

Figure 5 suggests that those who serve in TFA are more likely to be employed in a K - 12 school or in education more generally one to two years after their commitment ends. Our two stage least squares estimates suggest that serving in TFA increases the probability of being employed in a K - 12 school by 35.5 percentage points and in education more broadly by 47.5 percentage points. Individuals who serve are also 45.8 percentage points more likely to believe that service is
an important part of their career, and 35.9 percentage points more likely to prefer an urban teaching job over a suburban teaching job. Interestingly, there is not a statistically significant effect of service on wanting to work in education in the future, though the point estimate is positive. There is also no effect of service on the preference of an urban teaching job over a finance job at the same salary, though this may be because almost all survey respondents prefer teaching.

The effect of TFA on political idealism is depicted in Figure 6. Individuals were asked how liberal they consider themselves, and whether we should spend more money closing the achievement gap, on welfare assistance, and on fighting crime. Serving in TFA does not have a significant impact on political idealism, at least as we have measured it here. We cannot rule out moderate size effects of either direction, however.

Our final set of outcomes, racial tolerance, are presented in Figure 7. Our primary measure of racial tolerance comes from the brief format Implicit Association Test (IAT), developed by Sriram and Greenwald (2009), that assesses the relative strength of automatic associations between “good” words and black faces. We normalize the IAT to have a mean of zero and a standard deviation of one in each cohort. To complement the IAT measure, individuals were also asked about the relative intelligence of whites, Asians, blacks and Hispanics, and a set of six questions that together make up the Modern Racism Scale (McConahay 1983). We normalize the Modern Racism Scale to have a mean of zero and a standard deviation of one in each cohort. Each measure is normed so that higher numbers are associated with greater tolerance. More information on each measure is available in the data appendix.

Remarkably, serving in TFA increases implicit white-black tolerance by 0.980 standard deviations. To put this in context, black applicants score 0.558 standard deviations higher than Asian applicants on the IAT, while white and Hispanic applicants score 0.084 and 0.253 standard deviations higher than Asian applicants, respectively. The causal impact of TFA service is therefore equivalent to moving the median Asian applicant to above the 50th percentile of the black applicant distribution of racial tolerance towards blacks.

Serving in TFA is also associated with higher explicit white-black tolerance in the Modern Racism Scale, and a higher probability of believing that blacks and Hispanics are at least as intelligent as both whites and Asians, though none of the estimates for explicit racial tolerance are statistically significant. Taken literally, this implies that while there is little treatment effect on measures of explicit tolerance, TFA increases the unconscious tolerance of its members.

5.3 Analysis of Subsamples

Table 2 investigates heterogeneous treatment effects across gender, ethnicity, and whether or not a TFA applicant received a Pell Grant in college (a proxy for poverty). We allow for separate trends in interview score by group. The impact of service on faith in education and educational involvement is larger for men than for women. Service increases a male applicants faith in education by 2.261 standard deviations, while increasing a female applicant’s faith in education by 0.867 standard deviations. Educational involvement increases 3.289 standard deviations for male applicants and
0.917 standard deviations for female applicants.

There are no statistically significant differences by ethnicity, though the point estimates tend to be larger for blacks and Hispanics, even for the impact of service on racial tolerance. Estimated impacts are mixed by Pell Grant status, with Pell Grant recipients increasing their involvement in education more than non-recipients, but increasing their faith in education less than non-recipients.

5.4 Tests for Quasi-Random Assignment

Our empirical strategy assumes that applicants do not selectively respond to our survey. One specific concern is that former TFA corps members will be more likely to respond, or that there will be differences between the the alumni and non-alumni that respond to the survey. Such selective response could invalidate our empirical design by creating discontinuous differences in respondent characteristics around the score cutoff. Although the continuity assumption cannot be fully tested, we investigate whether the fraction of applicants responding changes at the cutoff, whether the observable characteristics of survey respondents trends smoothly through the cutoff, and examine the density of survey respondents around the cutoff.

Figure 1 tests whether the fraction of 2007 applicants responding to the survey changes at the cutoff. We present actual and fitted values of the survey response rate. There does not appear to be any difference in the response rate at the score cutoff. If anything, applicants above the cutoff are somewhat less likely to take to the survey.10

Figure 8 tests whether the observable characteristics of 2007 survey respondents trends smoothly through the cutoff. If there is a discontinuous change at the cutoff, that would indicate that respondents who were eligible for TFA differ in a way that would invalidate our research design. We present actual and fitted values for 2007 applicants to TFA who answered at least one question on our survey. Actual values are plotted in bins of size 0.0025. As with our first stage and reduced form regressions, the fitted values are from a regression of the dependent variable on an indicator variable for being above the cut-off score, a quadratic in interview score, and a quadratic in interview score interacted with the indicator for being above the cut-off score. We examine gender, ethnicity, college GPA, whether an individual had a Pell Grant in college, and whether an individual majored in a math or science field. Survey respondents with interview scores just above the cutoff have lower college GPAs, but are no more likely to be white or Asian, male, eligible for a Pell Grant, or to have majored in a math or science field. Results are identical in the full sample of applicants.11

10Appendix Figure 2 presents analogous results for the other TFA cohorts. For the 2007 - 2009 cohorts we control for interview score and interview score interacted with scoring above the cutoff. For the 2003 - 2006 cohorts we include fully nonparametric controls for each interview subscore, and report the point estimate for an indicator variable equal to one if the combination of interview subscores is eligible for TFA admission. While there is little difference in the response rates of eligible and ineligible applicants from the 2007 through 2009 cohorts, the marginal eligible applicant from the 2003 to 2006 is approximately 6 to 11 percentage points more likely to respond to our survey. The results from these cohorts should be interpreted with this caveat in mind.

11Appendix Figure 4 presents results from all cohorts, minus Pell Grant status which is not available from 2003 to 2006. We plot reduced form coefficients and associated 95 percent confidence intervals for each cohort, each from a separate regression. Eligible applicants who took the survey are more likely to be white or Asian in the 2003 through 2006 cohorts, and have somewhat higher college GPAs. Eligible and ineligible applicants do not differ by gender or
of the results from Figure 8 suggest the kind of selective survey response that would threaten our empirical design.

To further explore the potential impact of any selective survey response on our results, Appendix Table 8 presents summary statistics for all non-alumni members, non-alumni members who responded to our email strategies and took the survey in the first round, and non-alumni members who responded to our phone strategies and took the survey in the second round. There are no statistical differences between non-alumni members who responded to the email and phone recruitment strategies on the seven baseline variables available for all applicants.

As a final test of the potential impact of selective survey response on our results, Appendix Figure 5 presents our main results excluding non-alumni members who responded to our phone recruitment strategies. The results in Appendix Figure 5 are nearly identical to our earlier results, with the exception of racial tolerance, which is no longer statistically significant (p-value = 0.204). Taken together, we interpret the evidence in this section as suggesting that any potential bias from the differential recruitment strategies is likely to be minimal.

Our final robustness test is to check whether the frequency of respondents changes at the cutoff. To provide a formal estimate of a potential kink in the number of observations at the cutoff, we follow the approach of McCrary (2008) and first collapse the data into equal sized bins. The two key variables in the collapse data set are the number of observations in each bin and the interview score that each bin is centered around. We then regress the number of observations in each bin on a third order polynomial in interview score which we allow to vary on either side of the cutoff. A third order polynomial does a good job of fitting the data with an $R^2$ of 0.96 in the survey sample and 0.98 in the full sample. As suggested by Figure 9, the coefficient on the interaction term is statistically insignificant in both samples. Results are similar for both higher and lower order polynomials.

Given the general lack of statistical significance of our robustness checks, we interpret our results as showing no clear evidence that our identifying assumption is violated in our primary sample of 2007 applicants. Our robustness checks are less clear for the 2003 to 2006 cohorts. Eligible applicants are more likely to respond to our survey, and survey respondents are more likely to be white or Asian than non-respondents. The results from the 2003 to 2006 cohorts should be interpreted with this caveat in mind.

### 5.5 Additional Cohorts

One potential caveat to our analysis is that it includes only one cohort of TFA applicants surveyed roughly a year after their service commitment ended. If there are important longer term impacts of service, our analysis will understate the true impact of TFA. If, on the other hand, the impacts fade over time, our estimates are an upper bound on the true effects of TFA.

To shed some light on this issue, we also collected data on TFA applicants in the 2003 to 2006 and 2008 to 2009 cohorts. Recall that between 2003 and 2006, TFA admitted candidates who met prespecified interview score requirements. For example, TFA admitted all applicants with the college major in any of the cohorts in our sample.
highest possible critical thinking and organization skills conditional on having minimally acceptable scores in all other areas. Or, applicants could be admitted by having the highest possible score in perseverance and organizational ability, again conditional on minimally acceptable scores in all other areas. We estimate the impact of service in the 2003 to 2006 cohorts by instrumenting for TFA placement using an indicator variable equal to one if a candidate meets one of the six subscore criteria for admissions. We include fully nonparametric controls for each interview subscore. The impact of TFA service is therefore identified using the interaction of the subscores. Our key identifying assumption is that, conditional on our nonparametric controls, the interaction of interview subscores only impacts future outcomes through TFA placement. For candidates from the 2008 and 2009 cohorts, we use our regression discontinuity strategy outlined in Section 4.

Figure 10 presents results for the impact of service on our summary measures for all available cohorts. We plot reduced form coefficients and associated 95 percent confidence intervals for each cohort. Each estimate comes from a separate regression. The impact of service on educational and racial beliefs and educational involvement is persistent. Alumni from the 2003 to 2006 cohorts are more likely to believe in the power of education, more likely to be employed in education, and are more racially tolerant. Point estimates on the educational beliefs and involvement variables are statistically significant for all alumni cohorts. The racial tolerance point estimate is statistically significant at the 5 percent level for the 2003, 2005, and 2006 cohorts, and statistically significant at the 10 percent level for the 2004 cohort. On the other hand, current TFA corps members from the 2008 and 2009 cohorts are only somewhat more likely to believe in the power of education, and are no more racially tolerant than the marginal non-corps member.

6 Conclusion

Nearly one million American youth have participated in service programs such as Peace Corps and Teach For America, and annual government spending in support of youth service programs is many hundreds of millions of dollars. This paper has shown that serving in Teach For America has a positive impact on an individual’s faith in education, involvement in education, and racial tolerance. The impact of service is also quite persistent, with similar effects five years after the completion of the TFA service commitment. The impact of service on educational beliefs and involvement is larger for males, but there are no statistically significant differences by ethnicity.

Our results, particularly those on racial beliefs, are broadly consistent with the “Contact Hypothesis,” which suggests that contact with other groups will increase tolerance. Changes occur through a combination of increased learning, changed behavior, new affective ties, and reappraisals of one’s own group (Pettigrew 1998). A substantial empirical literature suggests that intergroup contact is negatively correlated with intergroup prejudice (Pettigrew and Tropp 2006). Recent research suggests that this correlation may be causal. Van Laar et al. (2005) and Boisjoly et al. (2006) show that white students at a large state university who were randomly assigned black roommates in their first year are more likely to endorse affirmative action, have more personal contact with
minority groups, and view a diverse student body as essential for a high-quality education. In a similar study, Barnhardt (2009) shows that Hindus randomly assigned to live near Muslim neighbors are significantly more implicitly and explicitly tolerant.

Recall, TFA service typically involves a considerable degree of intergroup contact over a two year period. Seventy-four percent of alumni members in our sample are white, and 80 percent have at least one parent with a college degree. The average parental income of a corps member is $118 thousand. In stark contrast, roughly 80 percent of the students taught by TFA members qualify for free or reduced-price lunch, and more than 90 percent of these students are African-American or Hispanic.

Taken together, the evidence presented in this paper suggests that TFA service has a significant impact on an individual’s values and career decisions. Youth service, particularly service involving extended periods of intergroup contact, may not only help disadvantaged communities, but help create a more socially conscious and more racially tolerant society.
References


Notes: This figure presents actual and fitted values for 2007 TFA applicants. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, interview score, and interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable.
Notes: This figure presents actual and fitted values for 2007 TFA applicants. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable.
Figure 3
Summary of Main Results

Notes: This figure presents actual and fitted values for 2007 TFA applicants who took the survey. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable. Each index was constructed by standardizing the sum of all questions in that area to have a mean of zero and a standard deviation of one. All standardized variables were converted to indicator variables using the median of the full sample. The variables included in each composite variable are available in the data Appendix.
Figure 4
Faith in Education

Notes: This figure presents actual and fitted values for 2007 TFA applicants. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable.
Figure 5
Involvement in Education

Notes: This figure presents actual and fitted values for 2007 TFA applicants. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable.
Figure 6
Political Idealism

Notes: This figure presents actual and fitted values for 2007 TFA applicants. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable.
Notes: This figure presents actual and fitted values for 2007 TFA applicants. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable.
Figure 8
Test of Quasi-Random Assignment

Notes: This figure presents actual and fitted values for 2007 TFA applicants who took the survey. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable.
Notes: This figure presents actual and fitted values for 2007 TFA applicants. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a cubic in interview score, and a cubic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable.
Figure 10
Main Results by Cohort

Faith in Education

Involvement in Education

Political Beliefs

Racial Tolerance

Notes: This figure presents point estimates and 95 percent confidence intervals for the reduced form effects by cohort. The 2007 - 2009 cohorts are estimate using a regression discontinuity design, control for a quadratic interview score and a quadratic in interview score interacted with scoring above the cutoff score. The 2003 - 2006 cohorts are estimate using the interaction between interview subscores that determines TFA selection. For the 2003 - 2006 cohorts we include fully nonparametric controls for each interview subscore, and report the point estimate for an indicator variable equal to one if the combination of interview subscores is eligible for TFA admission.
### Table 1
Summary Statistics

<table>
<thead>
<tr>
<th>Background Variables</th>
<th>TFA (1)</th>
<th>SD (2)</th>
<th>N (3)</th>
<th>Not TFA (4)</th>
<th>SD (5)</th>
<th>N (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>0.07</td>
<td>0.25</td>
<td>1,233</td>
<td>0.07</td>
<td>0.26</td>
<td>1,067</td>
</tr>
<tr>
<td>White</td>
<td>0.74</td>
<td>0.44</td>
<td>1,233</td>
<td>0.71</td>
<td>0.45</td>
<td>1,067</td>
</tr>
<tr>
<td>Black</td>
<td>0.06</td>
<td>0.23</td>
<td>1,233</td>
<td>0.07</td>
<td>0.25</td>
<td>1,067</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.07</td>
<td>0.25</td>
<td>1,233</td>
<td>0.06</td>
<td>0.24</td>
<td>1,067</td>
</tr>
<tr>
<td>Mixed Race</td>
<td>0.05</td>
<td>0.21</td>
<td>1,233</td>
<td>0.04</td>
<td>0.20</td>
<td>1,067</td>
</tr>
<tr>
<td>College GPA</td>
<td>3.58</td>
<td>0.29</td>
<td>1,233</td>
<td>3.46</td>
<td>0.35</td>
<td>1,067</td>
</tr>
<tr>
<td>Received Pell Grant</td>
<td>0.19</td>
<td>0.39</td>
<td>1,233</td>
<td>0.23</td>
<td>0.42</td>
<td>1,067</td>
</tr>
<tr>
<td>Math or Science Major</td>
<td>0.17</td>
<td>0.37</td>
<td>1,233</td>
<td>0.19</td>
<td>0.40</td>
<td>1,067</td>
</tr>
<tr>
<td>Married</td>
<td>0.12</td>
<td>0.33</td>
<td>1,233</td>
<td>0.13</td>
<td>0.34</td>
<td>1,067</td>
</tr>
<tr>
<td>Mother has BA</td>
<td>0.35</td>
<td>0.48</td>
<td>1,201</td>
<td>0.40</td>
<td>0.49</td>
<td>1,026</td>
</tr>
<tr>
<td>Father has BA</td>
<td>0.39</td>
<td>0.49</td>
<td>1,201</td>
<td>0.33</td>
<td>0.47</td>
<td>1,026</td>
</tr>
<tr>
<td>Father has more than BA</td>
<td>0.28</td>
<td>0.45</td>
<td>1,200</td>
<td>0.28</td>
<td>0.45</td>
<td>1,023</td>
</tr>
<tr>
<td>Father has more than BA</td>
<td>0.46</td>
<td>0.50</td>
<td>1,200</td>
<td>0.41</td>
<td>0.49</td>
<td>1,023</td>
</tr>
</tbody>
</table>

**Faith in Education**
- Poor children can compete with more advantaged children: 0.76, 0.43, 1,101 vs. 0.55, 0.50, 931
- The achievement gap is solvable: 0.57, 0.49, 1,100 vs. 0.41, 0.49, 932
- Fraction of minorities that should graduate college: 0.66, 0.26, 922 vs. 0.53, 0.27, 753
- Teachers are most important determinant of student success: 0.69, 0.46, 1,071 vs. 0.37, 0.48, 895
- Schools can close the achievement gap: 0.73, 0.44, 1,100 vs. 0.53, 0.50, 933
- Teachers can ensure most students achieve: 0.76, 0.43, 1,101 vs. 0.53, 0.50, 933

**Involvement in Education**
- Employed at K - 12 School: 0.45, 0.50, 1,233 vs. 0.20, 0.40, 1,067
- Employed in Education: 0.54, 0.50, 1,233 vs. 0.26, 0.44, 1,067
- Service Very Important: 0.80, 0.40, 1,146 vs. 0.72, 0.45, 971
- Prefer teaching over finance: 0.88, 0.32, 1,135 vs. 0.90, 0.31, 949
- Prefer urban school over suburban: 0.75, 0.43, 1,134 vs. 0.56, 0.50, 954
- Interested in working in education: 0.53, 0.50, 1,233 vs. 0.48, 0.50, 1,067

**Political Beliefs**
- Liberal: 0.66, 0.47, 1,092 vs. 0.64, 0.48, 923
- We should spend more closing the achievement gap: 0.88, 0.32, 1,041 vs. 0.85, 0.35, 875
- We should spend more on welfare assistance: 0.32, 0.47, 1,041 vs. 0.41, 0.49, 875
- We should spend more fighting crime: 0.38, 0.48, 1,041 vs. 0.44, 0.50, 875

**Racial Tolerance**
- IAT White-Black: 0.06, 1.02, 1,001 vs. -0.07, 0.97, 855
- Whites/Asians and Blacks/Hispanics are equally intelligent: 0.60, 0.49, 905 vs. 0.58, 0.49, 783
- White - Black Modern Racism Score: 0.09, 0.92, 946 vs. -0.10, 1.08, 794

*Notes: This table reports summary statistics for the 2007 TFA application cohort. The sample is all applicants who answered at least one survey question.*
### Table 2
#### Subsample Results

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>p-value</th>
<th>Asian/ White</th>
<th>Black/ Hispanic</th>
<th>p-value</th>
<th>Pell Grant</th>
<th>No Pell Grant</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Faith in Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td></td>
<td>2.261***</td>
<td>0.867***</td>
<td>0.006</td>
<td>0.892</td>
<td>1.142***</td>
<td>0.106</td>
<td>0.791***</td>
<td>1.286**</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.823)</td>
<td>(0.218)</td>
<td></td>
<td>(0.552)</td>
<td>(0.249)</td>
<td></td>
<td>(0.305)</td>
<td>(0.616)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2029</td>
<td>2032</td>
<td></td>
<td>1485</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Involvement in Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.289**</td>
<td>0.917***</td>
<td>0.023</td>
<td>1.413</td>
<td>1.471***</td>
<td>0.152</td>
<td>1.111**</td>
<td>−0.575</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>(1.444)</td>
<td>(0.364)</td>
<td></td>
<td>(0.985)</td>
<td>(0.410)</td>
<td></td>
<td>(0.509)</td>
<td>(0.976)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2290</td>
<td>2294</td>
<td></td>
<td>1667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Political Beliefs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.960</td>
<td>−0.068</td>
<td>0.225</td>
<td>−0.196</td>
<td>0.248</td>
<td>0.703</td>
<td>0.108</td>
<td>0.546</td>
<td>0.730</td>
</tr>
<tr>
<td></td>
<td>(0.790)</td>
<td>(0.215)</td>
<td></td>
<td>(0.516)</td>
<td>(0.253)</td>
<td></td>
<td>(0.311)</td>
<td>(0.594)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2019</td>
<td>2022</td>
<td></td>
<td>1480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Racial Tolerance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.738</td>
<td>0.648***</td>
<td>0.326</td>
<td>1.015</td>
<td>0.612**</td>
<td>0.142</td>
<td>0.406</td>
<td>1.421</td>
<td>0.226</td>
</tr>
<tr>
<td></td>
<td>(0.750)</td>
<td>(0.255)</td>
<td></td>
<td>(0.691)</td>
<td>(0.272)</td>
<td></td>
<td>(0.335)</td>
<td>(0.944)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1848</td>
<td>1852</td>
<td></td>
<td>1341</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports two-stage least squares estimates by subgroup. The sample is all 2007 applicants who answered at least one question included in the composite index. All regressions control for a quadratic in the interview score interacted with an indicator variable for scoring above the cutoff score. Standard errors are clustered at the interview score level. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
7 Online Appendix A: Data Description and Construction of Variables

NOT FOR PUBLICATION

Data for this project comes from a web-based survey administered between April 2010 and May 2011. This appendix describes these data and details the procedures used to code them.

7.1 Summary Indices

Racial Tolerance
This variable consists of the white - black IAT scores.

Faith in Education
This variable was constructed by standardizing the sum of our education belief questions to have a mean of zero and a standard deviation of one in the full sample. Rather than add dichotomous and standardized variables together, we converted all standardized variables to indicator variables. Specifically, we created an indicator variable equal to one if the variable was above the median of the full sample. The set of measures include whether poor children can compete with more advantaged children, whether the achievement gap is solvable, the percent of minorities that should be expected to graduate college, whether teachers are the most important determinant of student success, whether schools alone can close the achievement gap, and whether teachers can ensure most students achieve.

Involvement in Education
This variable was constructed by standardizing the sum of our employment questions to have a mean of zero and a standard deviation of one in the full sample. Rather than add dichotomous and standardized variables together, we converted all standardized variables to indicator variables. Specifically, we created an indicator variable equal to one if the variable was above the median of the full sample. The set of measures include whether an individual is employed at a K - 12 school, whether an individual is employed in education more broadly, whether an individual believes service is important, whether an individual is interested in working in education in the future, whether an individual prefers teaching in an urban district to a finance career, and whether an individual prefers teaching in an urban district to a suburban district.

Political Idealism
This variable was constructed by standardizing the sum of our political belief questions to have a mean of zero and a standard deviation of one in the full sample. Rather than add dichotomous and standardized variables together, we converted all standardized variables to indicator variables. Specifically, we created an indicator variable equal to one if the variable was above the median of the full sample. The set of measures include whether an individual self identifies as liberal or very liberal, whether an individual believes we should spend more on closing the achievement gap, whether an individual believes we should spend more on welfare, and whether an individual believes we should spend more on fighting crime.

7.2 Background

Parent’s Education
Respondents were asked "What is the highest level of education that your mother has completed?" and
"What is the highest level of education that your father has completed?" The answer choices range from less than a high school diploma to Ph.D., Ed.D., or D.B.A. We recoded this variable to be two variables. The first is equal to one if the respondent’s mother has a B.A. The second is equal to one if the respondent’s mother has more than a B.A. We recoded the father variable in the same way.

**Parent’s Income**
Respondents were asked "During your senior year of high school, what is your best estimate of your parents’ income?" The answer choices range from less than $20,000 to $100,000 in $20,000 bins, up to $150,000 in $25,000 bins and up to more than $250,000 in $50,000 bins. The responses to the earnings questions were transformed into real-valued variables using the mid-point of each bin and $275,000 for the top most bin.

**7.3 Racial Tolerance**

**Implicit Association Test**
We use a brief format Implicit Association Test (IAT), developed in Sriram and Greenwald (2009), to assess the relative strength of automatic associations between "good" words and black faces. The IAT relies on a respondent’s speed of response to measure the strength of their unconscious mental associations. The respondent must quickly categorize words and pictures of faces that appear on the screen. Faces are to be categorized as black or white and words as good or bad. Pairs of categories appear on either side of the screen. If the stimulus belongs to categories on the right (left) of the screen, respondents are to hit a key on the right (left) side of the keyboard. Each respondent completes a number of versions of the task. In the "compatible" versions, the two categories on one side are paired according to a stereotype, such as black with bad words and white with good words. In the incompatible versions, the categories are paired counter-stereotypically, such as black with good words and white with bad words. The key insight of the race IAT is that an implicit bias against blacks shows up as a response time differential between the compatible and incompatible versions.

We normalize the IAT measure so that it has a mean of zero and a standard deviation of one across the sample of survey respondents, with higher measures associated with a more anti-black response.

**Number of minority friends**
Respondents were asked "Of your 10 best friends, how many are black or Hispanic?" The variable was coded directly from the response.

**Minority relationship**
Respondents were asked "Have you ever dated someone that is black or Hispanic?" We recoded this variable to equal one if the respondent said yes.

**Blacks and Hispanics do not value education**
Respondents were asked if they agreed or disagreed with the statement "blacks and Hispanics do not value education to the same extent that whites and Asians do." We recoded this variable to equal one if the respondent did not disagree strongly.

**Blacks disadvantaged due to ability or will power**
Respondents were asked "On average, blacks have worse jobs, income and housing than whites. These differences are mostly due to..." The answer choices were discrimination, blacks being born with less ability to
learn, blacks not having the same opportunities for education, and blacks not having the same motivation or will-power. We recoded this variable to equal one if the respondent answered that blacks do not have the same will-power or ability.

Believes whites/Asians work harder than blacks/Hispanics
Respondents were asked "Where would you rank whites on this scale?" The answer choices range from almost all are lazy to almost all are hard working on a seven point scale. This question was repeated for Asians, blacks and Hispanics. We coded this variable to equal to one if respondents said that either whites or Asians were more intelligent than either blacks or Hispanics.

Believes whites/Asians are more intelligent than blacks/Hispanics
Respondents were asked "Where would you rank whites on this scale?" The answer choices range from almost all are unintelligent to almost all are intelligent on a seven point scale. This question was repeated for Asians, blacks and Hispanics. We coded this variable to equal to one if respondents said that either whites or Asians were more intelligent than either blacks or Hispanics.

Modern Racism Scale
Respondents were asked on a five point scale whether they agreed or disagreed with the following statements: 1) Over the past few years, blacks have gotten more economically than they deserve; 2) Over the past few years, the government and news media have shown more respect for blacks than they deserve; 3) It is easy to understand the anger of black people in America; 4) Discrimination against blacks is no longer a problem in the United States; 5) blacks are getting too demanding in their push for equal rights; 6) blacks should not push themselves where they are not wanted. Each item was rescaled so that higher numbers were associated with a more anti-black response, then a simple average was taken of the six questions. We then normalized the scale to have mean zero and standard deviation one across the sample of survey respondents.

7.4 Faith in Education

Poor children can compete
Respondents were asked if they agreed or disagreed with the statement "Students from low-income communities cannot be expected to do as well in school as students from more affluent communities." We recoded this variable to equal one if the respondent disagreed strongly.

Achievement gap is solvable
Respondents were asked if they agreed or disagreed with the statement "The student achievement gap between children in low-income and high-income areas is a solvable problem." We recoded this variable to equal one if the respondent agreed strongly.

Percent of minorities that should be expected to graduate from college
Respondents were asked "17 percent of blacks 25 and older currently have a college degree. What percent of minority children can we reasonably expect to graduate from a 4-year college?"

Teachers are the most important determinant of success
Respondents were asked "Who is most important in determining how well students perform in school?"
The available responses were students, teachers and parents. We recoded this variable to equal one if the respondent answered teachers.

**Schools alone can close the achievement gap**
Respondents were asked if they agreed or disagreed with the statement "Students can only succeed if they have parents or family at home helping them." We recoded this variable to equal one if the respondent strongly disagreed.

**Teachers can ensure most students achieve**
Respondents were asked if they agreed or disagreed with the statement "There really is very little a teacher can do to ensure that most of his/her students achieve at a high level." We recoded this variable to equal one if the respondent strongly disagreed.

**Teachers are the most important source of success**
Respondents were asked "Which do you believe is the most important source of student success?" The answer choices included student’s home background, student’s intellectual ability, student’s enthusiasm or perseverance, teacher’s attention to the unique interests and abilities of the student, teacher’s use of effective methods of teaching, and teacher’s enthusiasm or perseverance. We recoded the variable to equal one if the respondent chose any of the three sources related to teachers.

**Teachers can get through to all students**
Respondents were asked if they agreed or disagreed with the statement "If teachers try really hard, they can get through to even the most difficult or unmotivated students." We recoded this variable to equal one if the respondent agreed strongly.

**Teachers are responsible for keeping students in school**
Respondents were asked if they agreed or disagreed with the statement "I feel that teachers have the primary responsibility for keeping students from dropping out of school." We recoded this variable to equal one if the respondent agreed strongly.

**Teachers can ensure student success**
Respondents were asked if they agreed or disagreed with the statement "There really is very little a teacher can do to ensure that most of his/her students achieve at a high level." We recoded this variable to equal one if the respondent disagreed strongly.

**Students can succeed even without family support**
Respondents were asked if they agreed or disagreed with the statement "Students can only succeed if they have parents or family at home helping them." We recoded this variable to equal one if the respondent disagreed strongly.

**When a teacher should be fired**
Respondents were asked "A teacher should be dismissed if the following happens once." The answer choices included fails to make adequate progress on standardized tests, receives a poor evaluation from school administrators, receives a parent complaint, is found to have an inappropriate personal relationship with a student,
is found to have a past criminal record, commits a crime outside of school hours, makes an inappropriate racial/sexual remark to students in class, and physically disciplines a student. For each answer we recoded the variable to equal one if the respondent indicated yes.

### 7.5 Involvement in Education

#### Employed in a K - 12 School
Respondents were asked about their current employer. We coded a respondent as working at a K - 12 school if they reported working for a K-12 public school, a K-12 charter school, or a K-12 private school. Respondents were coded as not working in a K - 12 school if they worked for a for-profit company, a not-for-profit, a college or university, or a local, state or federal government.

#### Employed in K - 12 education
Respondents were asked about their current employer. We coded a respondent as working in education if they reported working for a K-12 public school, a K-12 charter school, a K-12 private school, or a not-for-profit that focuses on K-12 education. Respondents were coded as not working in K - 12 education if they worked for a for-profit company, a not-for-profit that does not focus on K-12 education, a college or university, or a local, state or federal government.

#### Prefers finance
Respondents were presented with the following scenario: "Consider two possible occupations: 1. Teaching in an urban public school district, 60 hours a week. 2. Working as a Vice President at a Middle Market Private Equity Firm, 60 hours a week. Both the teaching job and the finance job pay $50,000 per year. Which job would you prefer?" If the respondent chose teaching, the hypothetical salary for the private equity position was increased by $25,000. This scenario repeated itself until the respondent chose the finance job or the salary reached $500,000. If the respondent chose the private equity position, the hypothetical salary for the teaching position increased in the same manner. We recode this variable to equal one if the respondent prefers the teaching position at equal pay.

#### Prefers urban teaching
Respondents were presented with the following scenario: "Now consider two possible teaching positions: 1. Teaching in an urban public school district, 60 hours a week. 2. Teaching in a suburban public school district, 60 hours a week. Both the urban job and the suburban job pay $50,000 per year. Which job would you prefer?" If the respondent chose urban teaching, the hypothetical salary for the suburban teaching position was increased by $25,000. This scenario repeated itself until the respondent chose the suburban teaching job or the salary reached $500,000. If the respondent chose the suburban teaching position, the hypothetical salary for the urban teaching position increased in the same manner. We recode this variable to equal one if the respondent prefers the urban teaching position at equal pay.

#### Job satisfaction
Respondents were asked "Overall, how satisfied are you with this job?" The answer choices ranged from extremely satisfied to extremely unsatisfied on a six point scale. We recoded this variable to equal one if the respondent was extremely satisfied.

#### Importance of job characteristics
Respondents were asked "In any job, not just the one you have now, how important are each of the following aspects?" The characteristics included intellectual challenge, stress level, salary and benefits and service to society. The answer choices ranged from very important to not important on a three point scale. For each characteristic we recoded this variable to equal one if the respondent thought a characteristic was very important.

**Interested in future careers**

Respondents were asked "What is your level of interest in the following careers?" The careers included principal, elected office, political advocacy, business, law, science/technology, starting a social venture, work on Teach For America staff, and teaching. The answer choices ranged from high interest to low interest on a three point scale. For each career we recode this variable to equal one if the respondent has a high interest.

**7.6 Political Idealism**

**Liberal**

Respondents were asked "Where would you place yourself on this scale of political views?" The answer choices ranged from very liberal to very conservative on a five point scale. We recoded this variable to equal one if the respondent identifies as liberal or very liberal.

**Spending**

Respondents were asked "For each of the following issues, indicate if the government is spending too much money, the right amount, or too little." For each category we consider - lowering the crime rate, closing the achievement gap and increasing welfare/cash assistance for the poor - we recode this variable to equal one if the respondent believes we should spend more money on that issue.
8 Online Appendix B: Full TFA Applicant Web-Based Survey

8.1 Applicant Contact Process

Below, we detail the process used to contact TFA applicants regarding our survey.


TFA provided a total of 63,262 email addresses for applicants in the 2003 to 2009 cohorts. Each person received up to three emails between April 2010 to June 2010. A final email was sent to non-alumni in May 2011.

The fraction of emails that “bounced” was higher for non-alumni and older cohorts. 11 percent of 2003 alumni and 36 percent on non-alumni emails “bounced,” while only 2 percent of 2009 alumni and 6 percent on non-alumni emails “bounced.” Of those receiving the email, approximately 35 percent of alumni and 25 percent of non-alumni opened the email, with over 80 percent of those opening emails starting the survey.

B. Facebook

The second major touch point, for those who did not complete a survey after the initial emails, was to find TFA applicants on Facebook using accounts specifically created for the TFA/EdLabs survey. First, accounts for “Harvard EdLabs” were created on Facebook. Account pages included an overview of, and a link to, the survey. Second, we used email addresses from TFA to manually search for both alumni and non-admits on Facebook and attempted to “friend” them. Unfortunately, Facebook does not allow individual profiles to be used for any commercial gain, and 5 of the 16 accounts were disabled. Even for those accounts that were not disabled by Facebook, the outreach (i.e. the number of alumni and non-alumni that were successfully added as friends) was not particularly successful. Specifically, 2,612 friend requests were made resulting in 53 friend confirmations.

C. Phone calls (June 2010 – May 2011)

The third and final attempt at contacting TFA non-alumni was personal phone calls using phone numbers from TFA application records. Non-alumni that had not yet taken the survey were contacted via phone in the evenings during three phases. Starting in June 2011, 400 randomly selected numbers from the 2007 non-admits were called personally as part of a brief experiment. The 400 calls were further split into two groups. Two hundred randomly selected non-alumni members were offered a $20 Amazon gift card for completing the survey, while another randomly selected 200 were not offered any incentive. A final group consisting of the other 2,680 2007 non-admits were called using an automated call system. The 2,680 numbers that were called using an automated call system heard a brief, 30 second recording that provided context about and a link to the survey.

The two personal call groups produced statistically identical response rates, both of which were significantly higher than the automated call group. We therefore proceeded to call the remaining automated call group personally without offering incentives. During this second round of personal phone calls, 2,412 of the 2007 non-admits that were previously called using the automated phone system and had not yet taken the survey were called. As before, voicemails were left for those who did not answer the phone. In most cases, those people who were left messages were called again a few weeks later.

During the final third round, the non-admits from the 2003 - 2006 and 2008 - 2009 cohorts were called using the automated phone system then personal phone calls. These final calls all took place between April 2011 and May 2011.
8.2 Survey

Below is the full survey administered online between April 2010 and May 2011.

A. Demographics:

1. Please enter your preferred Email address below (e.g., jane.doe@acme.com).
2. Please enter your first name and last name below (e.g., John Smith).
3. Please indicate your birthdate using the dropdowns below.
4. Please indicate your sex.
5. Which of the following best describes your race/ethnicity? Asian, Pacific Islander, black, Non-Hispanic, black, Hispanic, Native American or Alaskan Native, white, Non-Hispanic, white, Hispanic, Mixed Race, black and white, Other Mixed Race, Other
6. What year did you apply to Teach For America?
7. Where do you live now? (e.g., 123 Single Street, Simpletown, WA, 92403)? Street Address, City, State, Zipcode
8. Taken all together, how would you say things are these days? Would you say that you are very happy, pretty happy or not too happy? Very happy? Pretty happy? Not too happy?

B. Background Information:

1. Where were you born? City, State, Country.
2. What High School did you graduate from? High School, City, State.
3. Which option below most accurately reflects your current relationship status? I am married, I am single (never married), I am living with someone in a marriage like relationship, I am separated, I am divorced, I am widowed.
4. How many children do you have?
5. What is the highest level of education that your mother has completed?, Less than a high school diploma, High School diploma, Some college/vocational school, Bachelor’s degree, Master’s degree, Law degree (JD, LLB), Medical degree (MD, DDS, DVM, etc.), Ph.D., Ed.D., D.B.A., Other/Not Applicable
6. What is the highest level of education that your father has completed? Same options as above.
7. During your senior year of high school, what is your best estimate of your parents’ income? Do not know, < $20,000, $20,000 to $39,999, $40,000 to $59,999, $60,000 to $79,999, $80,000 to $99,999, $100,000 to $124,999, $125,000 to $149,999, $150,000 to $199,999, $200,000 to $250,000, > $250,000
8. We are interested in your educational history SINCE you applied to Teach For America. Please fill in the level and type of degree for your three highest degrees obtained SINCE you applied to Teach For America (e.g., M.A. in Education).

C. Teach For America:
1. How likely is it that you would recommend Teach For America to a friend or family member? 
   Extremely Likely, Moderately Likely, Somewhat Likely, Slightly Likely, Not at all Likely
2. Did you serve in Teach For America?
3. Where did you serve? Type in School, City, State (e.g., South Eugene HS, Eugene, OR)
4. How many years did you teach in the district where you were placed?
5. Looking back, do you wish you had ...Taught in the district for more years, Taught in the district the 
   same amount of time, Taught in the district for fewer years
6. How satisfied were you with your principal at your placement school? 
   Extremely satisfied, Very satisfied, Satisfied, Unsatisfied, Very unsatisfied, Extremely unsatisfied
7. How satisfied were you with your relationship with other teachers at your placement school? Same 
   answer choices as above.
8. If you could do it all over again, would you serve as a Teach For America corps member?

D. Employment
1. Your current employer is a ...For-profit company, Not-for-profit that focuses on K-12 education, Not-
   for-profit that does not focus on K-12 education, K-12 public school, K-12 charter school, K-12 private 
   school, College or university, Local, state or federal government, Other
2. What is your current or most recent occupation? Please be as specific as possible (e.g., high school 
   math teacher)
3. What was your title when you started with your current employer? (e.g., Vice President of Sales)
4. What is your current title?
5. In what year did you start working for your current employer?
6. What is your current annual income? < $20,000, $20,000 to $39,999, $40,000 to $59,999, $60,000 
   to $79,999, $80,000 to $99,999, $100,000 to $124,999, $125,000 to $149,999, $150,000 to $199,999, 
   $200,000 to $250,000, > $250,000
7. What was your annual income when you started this job? Same as Above
8. How many hours per week do you typically work at this job? < 30 hours, 31-35 hours, 36-40 hours, 
   41-45 hours, 46-50 hours, 51-55 hours, 56-60 hours, 61-65 hours, 66-70 hours, > 70 hours
9. Overall, how satisfied are you with this job? Extremely Satisfied, Very Satisfied, Satisfied, Unsatisfied, 
   Very Unsatisfied, Extremely Unsatisfied
10. In any job, not just the one you have now, how important are each of the following aspects? Very 
    Important, Somewhat Important, Not Important
    - Intellectual Challenge
    - Stress Level
    - Salary and Benefits
    - Service to Society
11. In your current job, how satisfied are you with each of the following aspects? Very Satisfied, Somewhat Satisfied, Not Satisfied

   Same as above.

12. Consider two possible occupations:

   1. Teaching in an urban public school district, 60 hours a week.
   2. Working as a Vice President at a Middle Market Private Equity Firm, 60 hours a week.
   Both the teaching job and the finance job pay $50,000 per year.
   Which job would you prefer?

13. Now consider two possible teaching positions:

   1. Teaching in an urban public school district, 60 hours a week.
   2. Teaching in a suburban public school district, 60 hours a week.
   Both the urban job and the suburban job pay $50,000 per year.
   Which job would you prefer?

14. We are interested in your past jobs. What were your last three occupations and titles, prior to working for your current employer? Please be as specific as possible.

15. What is your level of interest in the following careers? High Interest, Some Interest, No Interest

   - Principal/Head of School
   - Elected Office
   - Political Advocacy
   - Business
   - Law
   - Science/Technology
   - Starting a Social Venture
   - Work on Teach For America staff
   - Teaching

E. Social and Civic Engagement:

1. In a typical month, how many total hours do you spend doing volunteer or charitable work? None, 1-5 hours, 5-10 hours, 11-15 hours, 16-20 hours, 21-25 hours, > 25 hours

2. Have you participated as a volunteer for any of the following groups during the past year? Select all that apply.

   - Educational Work with Kids (e.g. tutoring)
   - Other Educational Work (e.g. school board, school governance organizations)
   - Other Work with Kids (e.g. Big Brother/Big Sisters, coaching)
   - Other Volunteer Work (religious organizations, alumni organizations)
3. Have you donated money to any charitable organization or group during the past year? Which group? How much?
   Name of Organization: 
   Amount Donated: $

4. Of your 10 best friends, how many are black or Hispanic?

5. Have you ever dated someone that is black or Hispanic? Yes, No

F. Beliefs:

1. For each of the following issues, indicate if the government is spending too much money, the right amount, or too little.
   – Protecting the Environment
   – Improving the Nation’s Healthcare System
   – Lowering the Crime Rate
   – Reducing the level of Drug Addiction
   – Closing the Achievement Gap
   – Increasing Welfare/Cash Assistance for the Poor

2. Where would you place yourself on this scale of political views?
   Very Liberal
   Liberal
   Moderate
   Conservative
   Very Conservative

3. What type of school do you plan to send your children to?
   Traditional public school
   Charter school
   Magnet school
   Religiously affiliated or denominational private school
   Private prep school
   Not yet decided
   Not applicable as I am not planning to have children

4. Which do you believe is the most important source of student success?
   Student’s home background
   Student’s intellectual ability
   Student’s enthusiasm or perseverance
   Teacher’s attention to the unique interests and abilities of the student
   Teacher’s use of effective methods of teaching
   Teacher’s enthusiasm or perseverance
The next series of questions asks whether you agree or disagree with a particular statement.

5. The student achievement gap between children in low-income and high-income areas is a solvable problem. Agree strongly, Agree somewhat, Neither agree nor disagree, Disagree somewhat, Disagree strongly

6. If teachers try really hard, they can get through to even the most difficult or unmotivated students. Same as above.

7. I feel that teachers have the primary responsibility for keeping students from dropping out of school.

8. There really is very little a teacher can do to ensure that most of his/her students achieve at a high level.

9. Students can only succeed if they have parents or family at home helping them.

10. Students from low-income communities cannot be expected to do as well in school as students from more affluent communities.

11. Great schools can close the achievement gap.

12. What are the three factors you think are the most significant causes of the achievement gap?

13. 17 percent of blacks 25 and older currently have a college degree. What percent of minority children can we reasonably expect to graduate from a 4-year college?

14. Who is most important in determining how well students perform in school? Students, Teachers, Parents

15. A teacher should be dismissed if the following happens once. Select all that apply.
   Fails to make adequate progress on standardized tests
   Receives a poor evaluation from school administrators
   Receives a parent complaint
   Is found to have an inappropriate personal relationship with a student
   Is found to have a past criminal record
   Commits a crime outside of school hours
   Makes an inappropriate racial/sexual remark to students in class
   Physically disciplines a student

G. Knowledge:

1. On nationally standardized math and reading exams such as the Long Term National Assessment of Educational Progress (NAEP), the average black 8th grader tends to score how many grade levels behind the average white student? One grade level behind, Two grade levels behind, Three grade levels behind, Four grade levels behind, Five grade levels behind, Six or more grade levels behind

2. On nationally standardized math and reading exams such as the Long Term National Assessment of Educational Progress (NAEP), the average Hispanic 8th grader tends to score how many grade levels behind the average white student? Same as above.

3. What percentage of black, Hispanic, and white men aged 18 to 24 are incarcerated, on parole, or on probation? % of black men aged 18-24, % of Hispanic men aged 18-24, % of white men aged 18-24
4. What percentage of black, Hispanic, and white male youth are currently active gang members? % of black youth, % of Hispanic youth, % of white youth

5. What percentage of black, Hispanic, and white children are currently living in single parent households? % of black children, % of Hispanic children, % of white children

6. What percentage of black, Hispanic, and white children are born out of wedlock? Same as above.

7. What percentage of black, Hispanic, and white mothers are currently eligible for state or federal financial assistance (i.e. welfare)? % of black mothers, % of Hispanic mothers, % of white mothers

II. Beliefs

1. In the long run, hard work usually brings a better life and success; luck and connections don’t matter that much. Agree strongly, Agree somewhat, Neither agree nor disagree, Disagree somewhat, Disagree strongly

2. The government should take more responsibility to ensure that everyone is provided for. Same as above.

3. Poor people in this country can escape from poverty. Same as above.

4. Blacks and Hispanics do not value education to the same extent that whites and Asians do. Same as above.

5. Why are people poor in this country? They are poor because society treats them unfairly, They are poor because of laziness and lack of will power

6. On average, blacks have worse jobs, income and housing than whites. These differences are mostly due to ...Discrimination blacks being born with less ability to learn, blacks not having the same opportunities for education, blacks not having the same motivation or will-power

7. Where would you rank whites on this scale?
   Almost all are lazy
   Many more are lazy than hardworking
   More are lazy than hardworking
   Comparable numbers of lazy and hardworking
   More are hardworking than lazy
   Many more are hardworking than lazy
   Almost all are hardworking

8. Where would you rank blacks on this scale? Same as above.

9. Where would you rank Hispanics on this scale? Same as above.

10. Where would you rank Asians on this scale?

11. Where would you rank whites on this scale?
   Almost all are unintelligent
   Many more are unintelligent than intelligent
   More are unintelligent than intelligent
Comparable numbers of unintelligent and intelligent
More are intelligent than unintelligent
Many more are intelligent than unintelligent
Almost all are intelligent

12. Where would you rank blacks on this scale? Same as above.
13. Where would you rank Hispanics on this scale? Same as above.
14. Where would you rank Asians on this scale? Same as above.
15. Over the past few years, blacks have gotten more economically than they deserve. Agree strongly, Agree somewhat, Neither agree nor disagree, Disagree somewhat, Disagree strongly
16. Over the past few years, the government and news media have shown more respect for blacks than they deserve. Same as above.
17. It is easy to understand the anger of black people in America. Same as above.
18. Discrimination against blacks is no longer a problem in the United States. Same as above.
19. blacks are getting too demanding in their push for equal rights. Same as above.
20. blacks should not push themselves where they are not wanted. Same as above.
9 Online Appendix C: Additional Results

Appendix Figure 1
Response Rate by Cohort

Notes: This figure presents survey response rates for each TFA application cohort.
Notes: This figure presents point estimates and 95 percent confidence intervals for the reduced form difference in response rates by cohort. The 2007 - 2009 cohorts are estimate using a regression discontinuity design, control for a quadratic in interview score and a quadratic in interview score interacted with scoring above the cutoff score. The 2003 - 2006 cohorts are estimate using the interaction between interview subscores that determines TFA selection. For the 2003 - 2006 cohorts we include fully nonparametric controls for each interview subscore, and report the point estimate for an indicator variable equal to one if the combination of interview subscores is eligible for TFA admission.
Appendix Figure 3
Application Subscores in Full Sample

Notes: This figure presents actual and fitted values for all 2007 TFA applicants. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable. The dependent variable for each figure is the indicated interview subscore.
Appendix Figure 4
Test of Quasi-Random Assignment by Cohort

Notes: This figure presents point estimates and 95 percent confidence intervals for the reduced form effects by cohort. The 2007 - 2009 cohorts are estimate using a regression discontinuity design, control for a quadratic in interview score and a quadratic in interview score interacted with scoring above the cutoff score. The 2003 - 2006 cohorts are estimate using the interaction between interview subscores that determines TFA selection. For the 2003 - 2006 cohorts we include fully nonparametric controls for each interview subscore, and report the point estimate for an indicator variable equal to one if the combination of interview subscores is eligible for TFA admission.
Appendix Figure 5
Main Results in Email Only Sample

Faith in Education

Involvement in Education

Political Beliefs

Racial Tolerance

Notes: This figure presents actual and fitted values for 2007 TFA applicants who took the survey during the email recruitment phase. Non-alumni members who responded to the phone recruitment are excluded. The actual values are plotted in bins of size 0.0025. The fitted values come from a regression of the dependent variable on an indicator variable for scoring above the cutoff score, a quadratic in interview score, and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. The p-value is the significance of the indicator variable. Each index was constructed by standardizing the sum of all questions in that area to have a mean of zero and a standard deviation of one. All standardized variables were converted to indicator variables using the median of the full sample. The variables included in each composite variable are available in the data Appendix.
### Appendix Table 1
Robustness of Main Results to Bandwidth and Polynomial Choice

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faith in Education</td>
<td>1.462***</td>
<td>0.960***</td>
<td>0.871***</td>
<td>2.969**</td>
<td>1.568***</td>
<td>1.256***</td>
</tr>
<tr>
<td></td>
<td>(0.323)</td>
<td>(0.161)</td>
<td>(0.131)</td>
<td>(1.283)</td>
<td>(0.386)</td>
<td>(0.270)</td>
</tr>
<tr>
<td></td>
<td>1348</td>
<td>1939</td>
<td>2032</td>
<td>1348</td>
<td>1939</td>
<td>2032</td>
</tr>
<tr>
<td>Involvement in Education</td>
<td>1.898***</td>
<td>1.175***</td>
<td>1.057***</td>
<td>4.324*</td>
<td>1.915***</td>
<td>1.562***</td>
</tr>
<tr>
<td></td>
<td>(0.551)</td>
<td>(0.268)</td>
<td>(0.216)</td>
<td>(2.413)</td>
<td>(0.659)</td>
<td>(0.459)</td>
</tr>
<tr>
<td></td>
<td>1541</td>
<td>2193</td>
<td>2294</td>
<td>1541</td>
<td>2193</td>
<td>2294</td>
</tr>
<tr>
<td>Political Beliefs</td>
<td>0.408</td>
<td>0.123</td>
<td>0.083</td>
<td>0.771</td>
<td>0.465</td>
<td>0.258</td>
</tr>
<tr>
<td></td>
<td>(0.311)</td>
<td>(0.161)</td>
<td>(0.132)</td>
<td>(0.933)</td>
<td>(0.370)</td>
<td>(0.265)</td>
</tr>
<tr>
<td></td>
<td>1339</td>
<td>1929</td>
<td>2022</td>
<td>1339</td>
<td>1929</td>
<td>2022</td>
</tr>
<tr>
<td>Racial Tolerance</td>
<td>0.658*</td>
<td>0.552***</td>
<td>0.464***</td>
<td>0.527</td>
<td>0.763*</td>
<td>0.748***</td>
</tr>
<tr>
<td></td>
<td>(0.355)</td>
<td>(0.176)</td>
<td>(0.142)</td>
<td>(1.191)</td>
<td>(0.428)</td>
<td>(0.301)</td>
</tr>
<tr>
<td></td>
<td>1231</td>
<td>1766</td>
<td>1852</td>
<td>1231</td>
<td>1766</td>
<td>1852</td>
</tr>
<tr>
<td>Polynomial Order</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Notes: This table reports two-stage least squares estimates with various polynomials and bandwidths. The sample is all 2007 applicants who answered at least one question included in the composite index. All regressions control for a quadratic in interview score and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. Standard errors are clustered at the interview score level. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
### Appendix Table 2
#### Summary of Main Results

<table>
<thead>
<tr>
<th></th>
<th>FS</th>
<th>RF</th>
<th>TSLS</th>
<th>TSLS w/controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Faith in Education</td>
<td>0.313***</td>
<td>0.501***</td>
<td>1.598***</td>
<td>1.582***</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.129)</td>
<td>(0.442)</td>
<td>(0.439)</td>
</tr>
<tr>
<td>Involvement in Education</td>
<td>0.287***</td>
<td>0.554***</td>
<td>1.931***</td>
<td>1.880***</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.200)</td>
<td>(0.754)</td>
<td>(0.745)</td>
</tr>
<tr>
<td>Political Beliefs</td>
<td>0.309***</td>
<td>0.175</td>
<td>0.565</td>
<td>0.603</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.127)</td>
<td>(0.431)</td>
<td>(0.431)</td>
</tr>
<tr>
<td>Racial Tolerance</td>
<td>0.289***</td>
<td>0.270**</td>
<td>0.934*</td>
<td>0.980**</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.136)</td>
<td>(0.503)</td>
<td>(0.501)</td>
</tr>
</tbody>
</table>

Notes: This table reports first stage, reduced form, and two-stage least squares estimates. The sample is all 2007 applicants who answered at least one question included in the composite index. All regressions control for a quadratic in interview score and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. Standard errors are clustered at the interview score level. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
Appendix Table 3
Faith in Education

<table>
<thead>
<tr>
<th></th>
<th>FS</th>
<th>RF</th>
<th>TSLS</th>
<th>TSLS w/controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor children can compete with more advantaged children</td>
<td>0.312***</td>
<td>0.164***</td>
<td>0.526***</td>
<td>0.528***</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.062)</td>
<td>(0.201)</td>
<td>(0.201)</td>
</tr>
<tr>
<td></td>
<td>2028</td>
<td>2028</td>
<td>2028</td>
<td>2028</td>
</tr>
<tr>
<td>The achievement gap is solvable</td>
<td>0.312***</td>
<td>0.150**</td>
<td>0.482**</td>
<td>0.477**</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.065)</td>
<td>(0.213)</td>
<td>(0.212)</td>
</tr>
<tr>
<td></td>
<td>2028</td>
<td>2028</td>
<td>2028</td>
<td>2028</td>
</tr>
<tr>
<td>Fraction of minorities that should graduate college</td>
<td>0.320***</td>
<td>0.099***</td>
<td>0.310***</td>
<td>0.303***</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.038)</td>
<td>(0.122)</td>
<td>(0.121)</td>
</tr>
<tr>
<td></td>
<td>1672</td>
<td>1672</td>
<td>1672</td>
<td>1672</td>
</tr>
<tr>
<td>Teachers are most important determinant of student success</td>
<td>0.309***</td>
<td>0.148**</td>
<td>0.480**</td>
<td>0.480**</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.065)</td>
<td>(0.214)</td>
<td>(0.213)</td>
</tr>
<tr>
<td></td>
<td>1963</td>
<td>1963</td>
<td>1963</td>
<td>1963</td>
</tr>
<tr>
<td>Schools can close the achievement gap</td>
<td>0.312***</td>
<td>0.081</td>
<td>0.258</td>
<td>0.256</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.062)</td>
<td>(0.200)</td>
<td>(0.198)</td>
</tr>
<tr>
<td></td>
<td>2029</td>
<td>2029</td>
<td>2029</td>
<td>2029</td>
</tr>
<tr>
<td>Teachers can ensure most students achieve</td>
<td>0.313***</td>
<td>0.265***</td>
<td>0.848***</td>
<td>0.845***</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.060)</td>
<td>(0.223)</td>
<td>(0.221)</td>
</tr>
<tr>
<td></td>
<td>2030</td>
<td>2030</td>
<td>2030</td>
<td>2030</td>
</tr>
</tbody>
</table>

Notes: This table reports first stage, reduced form, and two-stage least squares estimates. The sample is all 2007 applicants who took our web-survey. All regressions control for a quadratic in interview score and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. Standard errors are clustered at the interview score level. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
### Appendix Table 4
Involvement in Education

<table>
<thead>
<tr>
<th></th>
<th>FS</th>
<th>RF</th>
<th>TSLS</th>
<th>TSLS w/controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed at K - 12 School</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>0.287***</td>
<td>0.100*</td>
<td>0.348*</td>
<td>0.355*</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.056)</td>
<td>(0.199)</td>
<td>(0.198)</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
</tr>
<tr>
<td>Employed in Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>0.287***</td>
<td>0.138**</td>
<td>0.481**</td>
<td>0.475**</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.059)</td>
<td>(0.211)</td>
<td>(0.210)</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
</tr>
<tr>
<td>Service Very Important</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>0.292***</td>
<td>0.130**</td>
<td>0.447**</td>
<td>0.458**</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.055)</td>
<td>(0.206)</td>
<td>(0.206)</td>
</tr>
<tr>
<td></td>
<td>2113</td>
<td>2113</td>
<td>2113</td>
<td>2113</td>
</tr>
<tr>
<td>Prefer teaching over finance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>0.309***</td>
<td>0.029</td>
<td>0.095</td>
<td>0.084</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.040)</td>
<td>(0.132)</td>
<td>(0.132)</td>
</tr>
<tr>
<td></td>
<td>2081</td>
<td>2081</td>
<td>2081</td>
<td>2081</td>
</tr>
<tr>
<td>Prefer urban school over suburban</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>0.286***</td>
<td>0.107*</td>
<td>0.375*</td>
<td>0.359*</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.060)</td>
<td>(0.214)</td>
<td>(0.212)</td>
</tr>
<tr>
<td></td>
<td>2085</td>
<td>2085</td>
<td>2085</td>
<td>2085</td>
</tr>
<tr>
<td>Interested in working in education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>0.287***</td>
<td>0.033</td>
<td>0.115</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.061)</td>
<td>(0.213)</td>
<td>(0.212)</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
</tr>
</tbody>
</table>

Notes: This table reports first stage, reduced form, and two-stage least squares estimates. The sample is all 2007 applicants who took our web-survey. All regressions control for a quadratic in interview score and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. Standard errors are clustered at the interview score level. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
Appendix Table 5
Political Idealism

<table>
<thead>
<tr>
<th></th>
<th>FS</th>
<th>RF</th>
<th>TSLS</th>
<th>TSLS w/controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Liberal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.310***</td>
<td>0.058</td>
<td>0.186</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.062)</td>
<td>(0.204)</td>
<td>(0.204)</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>2011</td>
<td>2011</td>
<td>2011</td>
</tr>
<tr>
<td>We should spend more closing the achievement gap</td>
<td>0.304***</td>
<td>0.040</td>
<td>0.131</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.044)</td>
<td>(0.148)</td>
<td>(0.148)</td>
</tr>
<tr>
<td></td>
<td>1912</td>
<td>1912</td>
<td>1912</td>
<td>1912</td>
</tr>
<tr>
<td>We should spend more on welfare assistance</td>
<td>0.304***</td>
<td>0.108*</td>
<td>0.356</td>
<td>0.373</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.065)</td>
<td>(0.234)</td>
<td>(0.234)</td>
</tr>
<tr>
<td></td>
<td>1912</td>
<td>1912</td>
<td>1912</td>
<td>1912</td>
</tr>
<tr>
<td>We should spend more fighting crime</td>
<td>0.304***</td>
<td>-0.009</td>
<td>-0.030</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.066)</td>
<td>(0.216)</td>
<td>(0.215)</td>
</tr>
<tr>
<td></td>
<td>1912</td>
<td>1912</td>
<td>1912</td>
<td>1912</td>
</tr>
</tbody>
</table>

Notes: This table reports first stage, reduced form, and two-stage least squares estimates. The sample is all 2007 applicants who took our web-survey. All regressions control for a quadratic in interview score and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. Standard errors are clustered at the interview score level. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
## Appendix Table 6
### Racial Tolerance

<table>
<thead>
<tr>
<th></th>
<th>FS</th>
<th>RF</th>
<th>TSLS</th>
<th>TSLS w/controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>Racial Tolerance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.289***</td>
<td>0.270**</td>
<td>0.934*</td>
<td>0.980***</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.136)</td>
<td>(0.503)</td>
<td>(0.501)</td>
</tr>
<tr>
<td></td>
<td>1852</td>
<td>1852</td>
<td>1852</td>
<td>1852</td>
</tr>
<tr>
<td><strong>Whites/Asians and Blacks/Hispanics are equally intelligent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.349***</td>
<td>0.056*</td>
<td>0.160*</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.070)</td>
<td>(0.203)</td>
<td>(0.201)</td>
</tr>
<tr>
<td></td>
<td>1685</td>
<td>1685</td>
<td>1685</td>
<td>1685</td>
</tr>
<tr>
<td><strong>White - Black Modern Racism Score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.305***</td>
<td>0.205*</td>
<td>0.674</td>
<td>0.633</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.143)</td>
<td>(0.490)</td>
<td>(0.490)</td>
</tr>
<tr>
<td></td>
<td>1738</td>
<td>1738</td>
<td>1738</td>
<td>1738</td>
</tr>
</tbody>
</table>

Notes: This table reports first stage, reduced form, and two-stage least squares estimates. The sample is all 2007 applicants who took our web-survey. All regressions control for a quadratic in interview score and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. Standard errors are clustered at the interview score level. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
### Appendix Table 7
Test of Quasi-Random Assignment

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>FS</th>
<th>RF</th>
<th>TSLS</th>
<th>TSLS full sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>White or Asian</td>
<td>0.287***</td>
<td>0.015</td>
<td>0.051</td>
<td>−0.032</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.059)</td>
<td>(0.206)</td>
<td>(0.116)</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
<td>7338</td>
</tr>
<tr>
<td>Black or Hispanic</td>
<td>0.287***</td>
<td>−0.013</td>
<td>−0.047</td>
<td>−0.016</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.033)</td>
<td>(0.118)</td>
<td>(0.085)</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
<td>7338</td>
</tr>
<tr>
<td>Male</td>
<td>0.288***</td>
<td>−0.069</td>
<td>−0.240</td>
<td>−0.240</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.055)</td>
<td>(0.201)</td>
<td>(0.201)</td>
</tr>
<tr>
<td></td>
<td>2290</td>
<td>2290</td>
<td>2290</td>
<td>2290</td>
</tr>
<tr>
<td>Received Pell Grant</td>
<td>0.287***</td>
<td>−0.045</td>
<td>−0.158</td>
<td>−0.080</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.049)</td>
<td>(0.173)</td>
<td>(0.099)</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
<td>7338</td>
</tr>
<tr>
<td>Math or Science Major</td>
<td>0.287***</td>
<td>0.004</td>
<td>0.015</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.045)</td>
<td>(0.158)</td>
<td>(0.087)</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
<td>7338</td>
</tr>
<tr>
<td>College GPA</td>
<td>0.287***</td>
<td>−0.069**</td>
<td>−0.240*</td>
<td>−0.227***</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.035)</td>
<td>(0.126)</td>
<td>(0.075)</td>
</tr>
<tr>
<td></td>
<td>2294</td>
<td>2294</td>
<td>2294</td>
<td>7338</td>
</tr>
</tbody>
</table>

Notes: This table reports first stage, reduced form, and two-stage least squares estimates for various baseline characteristics. The sample is all 2007 applicants. All regressions control for a quadratic in interview score and a quadratic in interview score interacted with an indicator variable for scoring above the cutoff score. Standard errors are clustered at the interview score level. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
## Appendix Table 8
**Summary Statistics by Survey Response**

<table>
<thead>
<tr>
<th></th>
<th>TFA</th>
<th>Not TFA</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Survey</td>
<td>Survey - All</td>
<td>All</td>
<td>1st Round Survey</td>
<td>2nd Round Survey</td>
<td>1st - All</td>
</tr>
<tr>
<td>White</td>
<td>0.772</td>
<td>0.812</td>
<td>0.059***</td>
<td>0.722</td>
<td>0.782</td>
<td>0.070***</td>
</tr>
<tr>
<td>Asian</td>
<td>0.055</td>
<td>0.061</td>
<td>0.009</td>
<td>0.070</td>
<td>0.073</td>
<td>0.063</td>
</tr>
<tr>
<td>Black</td>
<td>0.096</td>
<td>0.061</td>
<td>-0.052***</td>
<td>0.118</td>
<td>0.063</td>
<td>0.088</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.058</td>
<td>0.047</td>
<td>-0.017**</td>
<td>0.068</td>
<td>0.059</td>
<td>0.049</td>
</tr>
<tr>
<td>College GPA</td>
<td>3.567</td>
<td>3.579</td>
<td>0.017</td>
<td>3.414</td>
<td>3.463</td>
<td>3.457</td>
</tr>
<tr>
<td>Received Pell Grant</td>
<td>0.219</td>
<td>0.192</td>
<td>-0.040***</td>
<td>0.270</td>
<td>0.238</td>
<td>0.226</td>
</tr>
<tr>
<td>Math or Science Major</td>
<td>0.162</td>
<td>0.165</td>
<td>0.003</td>
<td>0.179</td>
<td>0.178</td>
<td>0.207</td>
</tr>
<tr>
<td>Observations</td>
<td>3767</td>
<td>1233</td>
<td></td>
<td>3601</td>
<td>478</td>
<td>589</td>
</tr>
</tbody>
</table>

Notes: This table reports summary statistics for the 2007 TFA application cohort. TFA alumni members were only recruited by email. First round non-alumni respondents were recruited by email. Second round non-alumni respondents were recruited by phone. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.
### Appendix Table 9

**Number of Observations**

<table>
<thead>
<tr>
<th>Bin Size</th>
<th>Survey Sample</th>
<th>Full Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0010</td>
<td>-0.319</td>
<td>1.598</td>
</tr>
<tr>
<td></td>
<td>(3.471)</td>
<td>(7.056)</td>
</tr>
<tr>
<td></td>
<td>112</td>
<td>118</td>
</tr>
<tr>
<td>0.0025</td>
<td>-2.440</td>
<td>1.694</td>
</tr>
<tr>
<td></td>
<td>(8.962)</td>
<td>(24.762)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>0.0050</td>
<td>-3.230</td>
<td>18.846</td>
</tr>
<tr>
<td></td>
<td>(21.866)</td>
<td>(47.167)</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>0.0075</td>
<td>-11.581</td>
<td>59.263</td>
</tr>
<tr>
<td></td>
<td>(20.629)</td>
<td>(59.457)</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>0.0100</td>
<td>-71.076**</td>
<td>-61.337</td>
</tr>
<tr>
<td></td>
<td>(26.195)</td>
<td>(37.426)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

Notes: This table reports reduced form results testing for differences in the number of observations around the interview score cutoff. The sample is 2007 TFA applicants. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.