

# Who Benefits Most from a Same-Race mentor? Optimal Matching in a Nationwide Youth Mentoring Program \*

Brachel Champion,<sup>†</sup> Zachary Szlendak,<sup>‡</sup> and Corey Woodruff<sup>§</sup>

JULY 9, 2021

## ABSTRACT

We identify the impacts of assigning a mentor of the same race or ethnicity on the social, emotional and academic development of youth relative to assigning a mentor of a different race or ethnicity. Using the universe of matches from a nationwide youth mentoring program, we estimate a selection-on-observables model that is motivated by the matching heuristic match specialists use and the balance of pre-match observables across same-race status. We find that on average youth in same-race/ethnicity mentoring relationships had no significant improvements compared to those in cross-race matches after a year of mentoring. However, Black and Hispanic youth who were assigned a same-race/ethnicity mentor had slightly higher self-perceived school ability and improved attitudes concerning risky behaviors. In contrast to previous work, we find little evidence that race/ethnicity-matching improves grades or expectations for future educational attainment. These results imply that when minority mentors are in short supply, matching on race or ethnicity at the expense of another desirable trait may not lead to improved youth development.

*JEL:* H51, I13, J26

*Keywords:* youth mentoring, race congruence, matching, child development

---

\*We are grateful to Francisca Antman, Tania Barham, Kyle Butts, Brian Cadena, Hannah Denker, Richard Mansfield, Brian Marein, Terra McKinnish, Bernadette Sanchez, Evelyn Skoy, Heather Taussig and participants of the CU Boulder Labor Economics Seminar for helpful comments. We are indebted to the national agency which generously provide the data, and to administrators of the Denver, Colorado office who walked us through the matching process. The views expressed by the authors do not necessarily reflect the views of the organization in question, the Institute for Defense Analyses, or Amazon, Inc.

<sup>†</sup>Corresponding author: Department of Economics, University of Colorado Boulder, 256 UCB, Boulder, CO 80309; [brachel.champion@colorado.edu](mailto:brachel.champion@colorado.edu)

<sup>‡</sup>Institute for Defense Analyses; [zszlenda@ida.org](mailto:zszlenda@ida.org)

<sup>§</sup>Amazon, Inc; [corey.woodruff@colorado.edu](mailto:corey.woodruff@colorado.edu)

## 1 — Introduction

In 2014, an estimated 4.5 million youth in the U.S. were in structured, one-to-one mentoring relationships (Bruce & Bridgeland, 2014). Youth mentoring programs are held in high regard for their positive impacts on participant development. Experimental evaluations of mentoring programs have shown that mentored youth have better social-emotional skills (Grossman & Tierney, 1998) and increased likelihood of completing high school and enrolling in post-secondary education relative to non-mentored youth (Falk, Kosse, & Pinger, 2020; Rodriguez-Planas, 2012).<sup>1</sup> A shared racial or ethnic identity between the mentor and mentee is believed to be a primary determinant of the growth in youth outcomes, and many mentoring organizations prioritize race-congruency in the matching process. While the benefits of race-matching has been well-documented in a classroom context (Dee, 2004; Egalite, Kisida, & Winters, 2015; Harbatkin, 2021), the effects of assigning a same-race or ethnicity mentor on youth outcomes is less clear.

To answer this question, we estimate the impact of assigning a same-race/ethnicity mentor on the social, emotional and academic outcomes of the universe of youth participating in a large, nationally available mentoring program. The program supported over 135,000 mentoring relationships in 2019 across 200 local agencies in all 50 states. Administrators and staff often prioritize racial or ethnic matching in selecting mentors for youth. Furthermore, many of the youth participants and their parent(s) express preferences for a mentor of the same race or ethnicity. But there is often a mismatch between the supply of minority youth and minority mentors. In 2018, 72% of youth in the program in question were minorities, while only 32% of volunteers were non-white.<sup>2</sup> As a result, youth mentoring organizations must choose how to allocate a scarce supply of minority mentors to mentees. Therefore, it is important to understand which outcomes are most improved by race/ethnicity-matching, and for which youth this premium is largest.

It is not clear ex-ante that same-race matching should be preferred to cross-race. Rhodes, Reddy, Grossman, and Maxine Lee (2002) provides a detailed explanation of the potential social and cultural costs and benefits of both same- and cross-race mentor matching. The authors hypothesize that having a mentor who shares a racial or ethnic background can promote trust and aid in establishing a relationship. Mentors of a different racial or ethnic background may not

---

<sup>1</sup> DuBois, Portillo, Rhodes, Silverthorn, and Valentine (2011) provide a thorough review of the literature evaluating youth mentoring programs.

<sup>2</sup> Based on authors' calculations.

be able to adequately support the youth in the challenges they face. Successful role models of similar racial background may also empower youth—especially of minority heritage—to achieve higher levels of success themselves. On the other hand, cross-race matches can lead to many of the same positive outcomes as same-race matches. Racial diversity in mentoring may bridge cultural and social gaps and challenge cultural beliefs, or foster a sense of community between racial groups. Furthermore, when same-race mentors are in short supply, a cross-race match may be better for the youth than no match at all.

Indeed, the literature on race-matching in youth mentorship has found a range of impacts. In support of same-race matches, Rhodes et al. (2002) find that minority boys had declines in scholastic competence and self-worth relative to non-mentored minority boys, but that same-race matches experienced less of a decline than cross-race matches. Minority girls in same-race matches likewise reported slower declines in school value and self-worth than cross-race matched minority girls. However, the authors also found several benefits to cross-race matches. Minority youth were less likely to report initiating alcohol when placed in cross-race matches. In addition, youth in cross-race relationships reported that they were more likely to talk to their mentors when distressed and were more likely to describe their mentor as providing unconditional support. Finally, parents of youth in cross-race matches were more likely to believe that the relationship improved their children's peer relationships, that the mentor tried to build on the youth's strengths, and that the mentors took them to places they wanted to go.

In our study, the observed variation in race matching among completed matches comes from two sources. First, there is variation across local agencies in the relative supply of minority mentors. This variation is in part generated by neighborhood characteristics that could directly influence youth outcomes. Second, there is race-matching variation within an agency that arises from match specialists curating matches. Our discussions with administrators and staff lead us to conclude that matches are made largely on the basis of race, gender, shared interests and geographic proximity between the mentor and the youth, conditional on the availability of an eligible mentor. Ex-ante, there is no reason to prefer either source of variation. The cross-agency variation is not subject to idiosyncratic match specialist selection bias, but may be driven primarily by local economic and social factors, such as economic activity and crime, that are correlated with youth outcomes. On the other hand, within-agency variation generated by match specialists is likely uncorrelated with local characteristics but may be correlated with unobserved youth characteristics. Section 2 outlines the within-agency matching process that leads to our

preferred match-level, selection-on-observables model. The within-agency variation in race-matching is plausibly exogenous conditional on observed match characteristics because match specialists must balance all the dimensions of potential matches when selecting a mentor to assign. Key to our analysis, we use the rich set of controls available in the data to show the pre-match characteristics of the youth and mentors as well as youth's baseline outcomes do not differ substantially between same- and cross-race matches. However, if unobserved mentor or youth characteristics affect both the probability of same-race assignment and growth in the youth's outcomes, then estimates of same-race effects would be biased. To address match-level unobservables, we estimate a similar model at the agency level that removes bias from endogenous match formation, but at the cost of potentially exacerbating bias from local effects. Results are similar across both models which suggests that bias from either source of variation is likely minimal.

We use the universe of matches facilitated by the program from 2010 to 2018 to estimate the causal effect of assigning a same-race/ethnicity mentor (hereafter referred to as “same-race” for brevity) on the youth's social, emotional and academic outcomes after a year of mentoring. On average, we find that youth matched to a same-race mentor do not have significantly higher growth in outcomes relative to those in cross-race matches. However, minority youth experience some improvements when paired with a mentor of the same race: Black and Hispanic youth in same-race matches experience small but statistically significant improvements in their risk attitudes relative to their cross-race counterparts, and Black youth had higher self-perceived school ability. On the other hand, Hispanic youth in cross-race matches had higher grades in reading, social studies and science, and Black youth were more likely to report having a non-guardian mentor in their lives after a year of cross-race mentoring. Our findings support the theory that having a same-race mentor improves the self-esteem and confidence of certain youth, perhaps by sharing a background that creates a stronger relationship. But there are also benefits to cross-race mentoring that may complement the areas where same-race mentoring does not seem to impact the youth, such as course grades. These results imply that race-matching is an important dimension for youth mentoring organizations to consider when targeting the social and emotional development of minority youth. However, when the supply of minority mentors is scarce, matching on race at the expense of other important traits may not produce the fastest growth in youth. It may even be the case that, for some youth, cross-race mentoring has even higher benefits.

We add to the literature on youth mentoring by estimating the effect of race-matching on the outcomes of all youth participating in the program during 2010-2018. Grossman and Tierney (1998) and Herrera, Grossman, Kauh, and McMaken (2011) randomly assigned youth to receive mentorship from the Big Brothers Big Sisters (BBBS) program and found mentored youth had improved self-perceptions of scholastic ability, risk attitudes, and self-esteem relative to youth in the comparison group. In both studies, the sample sizes of about 1,100 youth are small relative to overall scale of BBBS and neither examine the impacts by the type of mentor assigned. Rhodes et al. (2002) use the experimental variation in Grossman and Tierney (1998) to estimate the effect of race-matching, but the race of the mentor was not randomized among the treated group. Hence, it is not clear if same-race mentoring increases youth development at a faster rate than cross-race mentoring or improves outcomes relative to non-mentored youth. We extend this literature by examining the universe of youth participants in the program and provide quasi-experimental evidence of the impacts of race-matching that accounts for the non-random assignment of minority mentors.

We also contribute to the literature on racial congruence by isolating the so-called “passive effects” of an adult mentor on the youth’s self-perceived social, emotional and academic abilities. Much of the existing work is focused on educational contexts, particularly K-12 schooling, where student outcomes are affected through “active teacher effects” and “passive teacher effects” as described in Harbatkin (2021). In the first case, teachers may leverage their authority—either consciously or unconsciously—to assign lower grades or dole out harsher punishments for misbehavior to students of a different race (Bates & Glick, 2013; Dee, 2005; Ouazad, 2014). At the same time, teachers may have a passive influence on the student by appealing to their motivations and self-confidence in daily interactions, which may have differential impacts by racial congruence (Van Ewijk, 2011). In the youth mentoring context at hand, the adult mentors have little to no formal authority or evaluative power over the youth, allowing us to attribute the estimated effects to the passive channel. In contrast, studies that examine the effect of same-race teacher assignment on test scores (Dee, 2004; Egalite et al., 2015; Harbatkin, 2021) or career path choice (Kofoed & McGovney, 2019) cannot disentangle the passive effects from the active effects of race-matching.

The rest of the paper is structured as follows. Section 2 outlines the youth mentoring program, the process of matching mentor to youth and the real-world constraints that produce variation in race matching that is plausibly exogenous after conditioning on a set of observed covariates.

Section 3 describes the data on the universe of matches during the 2010-2018 period used. Section 4 formalizes the empirical strategies used to identify the causal effect of race-matching on youth outcomes. Section 5 displays the results of estimating these models and discuss their implications. Finally, Section 6 concludes.

## **2 — Background**

The organization in question operates a volunteer-based, one-on-one youth mentoring program that pairs a youth mentee and an adult mentor. With over 200 local agencies across all 50 states, the program is accessible in most areas and most youth who apply are eligible to participate. This program supported over 135,000 matches in 2019, making it one of the largest youth mentoring programs in the U.S. Within the organization, matches are either “site-based” (SB) or “community-based” (CB) mentoring. SB matches are typically organized at a specific location, such as a workplace, school, or community center and tend to have more structure imposed on them by the local agency. In contrast, CB matches spend time in their community in activities like playing games at a park, attending a sporting event, or visiting a museum. mentors are expected to become a role model for their youth by consistently spending time together (e.g. 3-4 times a month). mentors in CB matches are encouraged to plan activities that foster a friendship with their youth, such as going for a walk or playing a board game. The program does not expect mentors to invest large amounts of money in their youth or spend time tutoring. Rather, the goal for the relationship is to inspire the youth through positive interactions and “quality time.”

The matching process begins with potential mentors and youth applying for the program at their local agency. youth are typically between the ages of 8 and 13 when entering the program, and can stay enrolled through age 18. youth (or their guardians) submit an application that includes basic demographic information as well as information on their preferences for a mentor. Once they have applied, an assigned match specialist, referred to as a “match specialist”, administers a baseline survey. mentors are usually 21 years or older, must pass a background check, and must complete an in-person interview with their match specialist. After the interview is complete, the adult enters the pool of available mentors. mentors are assigned to a youth by the match specialist based on a variety of criteria that can include gender, race, shared interests or background, travel time between the two, and the preferences stated by either party. Once assigned, the mentor and youth meet and mutually agree to form the match. If the match is

successful, the local agency administers a follow-up survey every twelve months the match survives. Either party may end the match at any time, though the match specialist provides continual support to avoid early terminations. In our sample (described in detail in Section 3), matches typically last for 2-3 years (mean: 34.93 months, median: 29.20 months).

It is common for youth and their parent(s) to state a preference for same-gender matches, and many request a mentor of the same race.<sup>3</sup> However, the youth's preferences are constrained by the available supply of mentors of a given race or gender. Figure 1 shows a majority of mentors in our sample of matches are white (69%), while a majority of youth are Black or Hispanic (66%); though almost all matches are same-gender. In addition to balancing preferences, match specialists face a limited supply of non-white mentors. Figure 2 shows the fraction of mentors and youth participating in a match within a local agency, separated by race. A point above the plotted 45 degree line implies mentors of that race are relatively overrepresented at their agency compared to youth, while a point below the diagonal line implies underrepresentation and a point on the line shows that the proportions of mentors and youth of a particular race are balanced. Figure 2 shows that in our sample, often a majority share of mentors are white while a majority share of youth are Black or Hispanic within a given agency.<sup>4</sup>

### 3 — Data

Our data contain the universe of matches between 2010 and 2018. The data include information on the race, gender and age of every mentor and youth that participated in BBBS mentorship in the United States. Also included are measures of the youth's socioeconomic status (number and type of guardians in the household, on free/reduced lunch) and the mentor's educational attainment. Most importantly, our data set also includes a survey which is administered to every youth upon entering the program (hereafter referred to as the baseline survey) and then again every year (12 months or school year for community- and school-based<sup>5</sup> matches, respectively) that the match continues (hereafter referred to as the follow-up survey). The survey includes 33

---

<sup>3</sup> We assume youth reveal their true preferences since requesting a mentor with certain qualities is costless to the youth and does not affect the probability of being matched overall.

<sup>4</sup> Anecdotally, the director of the local agency informed us that the supply of mentors is often mismatched with the supply of youth: most mentor applicants are white and/or female and most youth applicants are minorities and/or male.

<sup>5</sup> School-based matches are a subset of site-based matches.

Likert scale questions regarding the youth's outlook on their school experience, education and social experience. These 33 questions are aggregated into eight summary scores generated by taking the average of the component scores: social acceptance, school ability, truancy, grades, education expectations, risk attitudes, parental trust, and special adult.<sup>6</sup>

The school experience group is comprised of social acceptance, school ability, and truancy. The social acceptance score is derived using the social competence subscale from the Perceived Competence Scale for Children (PCSC, Harter (1982)) which measures the youth's perceptions of their friendships (e.g. "I have a lot of friends."). The school ability score is a shortened version of the cognitive competence module from the PCSC and measures the youth's self-perceived academic ability (e.g. "I am very good at my schoolwork", and "I feel that I am just as smart as other kids"). Truancy is the average days late to school and absent from the youth's self-reports.<sup>7</sup>

The education group contains two outcomes: grades and education expectations. Grades is the average letter grade (mapped from F-A to 1-5 correspondingly) the youth received in mathematics, reading or language arts, social studies, and science. Education expectations contains three questions regarding their prospective educational attainment: how likely they are to (1) finish high school, (2) go to college, and (3) finish college.

Lastly, the social experience group includes risk attitudes, parental trust, and special adult. Risk attitudes was adapted from the Peer Pressure Inventory (PPI) developed by Brown, Clasen, and Eicher (1986) and measures the youth's perceptions of whether certain risky behaviors are acceptable among kids their age. These behaviors include using tobacco, drugs and alcohol, truancy, and misbehavior (e.g. hitting someone and breaking rules in school). Parental trust is an abbreviated version of the Inventory of Peer and Parental Attachment (IPPA, Armsden and Greenberg (1987)) questionnaire which measures the youth's perceptions of their relationship with their parents (e.g. "My parents respect my feelings."). Lastly, special adult is a single dummy indicating whether the youth feels that they have a non-guardian adult in their life who is a role model to them.

In almost all cases, youth in both same- and cross-race matches experience growth in their outcomes on average. Table A1 shows the means of the eight outcomes were higher at first follow-up than their baseline values for both groups with one exception: grades were slightly

---

<sup>6</sup> The survey also includes questions on juvenile arrest but we omit these outcomes due to lack of variation—most youth in our sample were never arrested.

<sup>7</sup> Frequencies are categorically binned 1-4, i.e. 1=no absences, 2=1-2 absences, etc.



lower after a year of mentoring, though the difference is not statistically significant (results not shown). Of particular importance, the proportion of youth who identified a special adult in their life at follow up was approximately thirty percentage points higher than the baseline fraction. The definition of special adult describes a non-guardian mentor with whom the youth has a close connection, suggesting that the program is highly effective in assigning mentors who are able to bond with their youth. Although we cannot empirically test the growth in these outcomes against a counterfactual youth who did not receive mentoring, these changes over time suggest that mentoring improves the lives of participants after only twelve months.

We restrict our sample to the set of successful CB matches that lasted at least one year. Our analysis examines the change in youth outcomes at the first follow-up survey relative to their baseline survey. Hence, we omit any matches that do not have a baseline survey. This excludes all matches that began prior to 2010 since the match specialists did not begin administering a baseline survey until 2010. We exclude any matches that do not have at least one follow-up survey. It is possible that matches that did not last through the first follow-up survey differ in some important ways from those that did last. Table A2 tests for differences between matches that lasted through the first follow-up and those that did not. Imbens and Wooldridge (2009) recommend using a normalized difference in means test to account for statistically significant differences that arise simply due to large sample sizes. The authors suggest that normalized differences of less than 0.25 in absolute value indicate no significant difference between the two groups. We do not find evidence that these matches differ in any important ways based on observable characteristics of either the mentor or the youth. Lastly, we consider only CB matches because the matching process for SB programs tend to vary more by state, SB matches tend to end earlier, and youth enrolled in SB programs were not balanced across same-race status. The resulting sample includes about 29,532 matches. We do not observe any information on proposed matches that failed to form. Many agencies operate a waiting list for youth who were not able to be matched, so we assume that the probability of a youth rejecting a proposed match is low.

## **4 — Methodology**

Because mentors are assigned to youth by a match specialist, the variation in same-race status is non-random. However, agencies face shortages in the supply of minority mentor applicants which restricts the specialist's ability to be overly selective when assigning a mentor to the youth.

Figure 2 shows the fraction of mentors and youth at a local agency that are in each race category. In our sample, a majority of mentors at an agency are often white while a majority of youth are often Black or Hispanic. This within-sample stylized fact coincides with the experience of staff we spoke to: the pool of mentor applicants is predominantly white while the pool of youth applicants is predominantly racial or ethnic minorities.<sup>8</sup> As mentioned previously, families often express preferences for a mentor who is the same gender and race as the youth. Additionally, they commonly ask to be matched with a mentor that shares interests or has a particular level of education. These requirements further restrict the pool of eligible mentors that fulfill the requests of the youth. Our data support this conclusion. Figure 1 shows that almost all matches are same-gender but there is significant variation in same-race status. Match specialists face the difficult challenge of balancing the stated preferences of the youth with other dimensions of the match that increase the expected match length such as geographic proximity.

The constraints faced by match specialists when choosing a mentor to assign a youth motivate a selection-on-observables empirical strategy. Conditional on the race of the youth, the probability of matching a same-race and otherwise-eligible mentor depends on the local supply of applicants. If the pool of eligible minority mentor applicants is large, the probability of being assigned a same-race mentor is high and case-workers would have the freedom to match on other unobservable characteristics, potentially biasing our results. This does not seem to be the case (see Figure 2), however we only observe the mentors who were matched which represents a subset of the pool of applicants. If the local demographic composition of potential mentors is endogenous to the growth in youth's outcomes, or if the probability of an eligible mentor applying to the program is correlated with the observed characteristics of youth in the area, our results would be biased. Table 1 shows that same- and cross-race matches are observably similar at baseline which suggests there is youth selection bias induced by the match specialist. However, we cannot directly test for selection on unobservable match characteristics. To combat this possibility, we estimate specifications at both the match- and agency-level. After controlling for demographic and socioeconomic characteristics of the mentor and youth in each, the remaining variation in same-race status is plausibly exogenous, though each approach is susceptible to different types of bias. We describe the assumptions behind both models and the associated threats to identification below.

---

<sup>8</sup> Specifically, our contact at the local agency estimated that about 80% of mentors applicants were white, while 80% of youth applicants were racial or ethnic minorities.

#### 4.1. Match-Level Estimation

We first identify the impacts of assigning a same-race mentor on the outcomes of the youth at the match-level. Specifically, we estimate the equation

$$Y_{iat}^F = \alpha + \beta \text{SameRace}_{iat} + \eta Y_{iat}^B + \Gamma' \text{YouthChars}_{iat} + \Theta' \text{MentorChars}_{iat} + \delta_t + \varepsilon_{iat} \quad (1)$$

for match  $i$  at agency  $a$  in year  $t$ .  $\text{SameRace}_{iat}$  equals one if the mentor and the youth in match  $i$  are of the same race,<sup>9</sup> and  $Y_{iat}^F$  denotes the youth's outcome at the first follow-up survey. We control for the initial value of the outcome variable,  $Y_{iat}^B$ , to account for any baseline differences among youth.  $\text{YouthChars}_{iat}$  is a vector of the youth's race, age, gender, free/reduced school lunch status, and single parent home status.  $\text{MentorChars}_{iat}$  is a vector of the mentor's race, age, gender, and educational attainment.  $\delta_t$  are year fixed effects that account for any trends in outcomes over time. We cluster the standard errors at the agency level to allow for dependency in the error term among matches within an agency.

To interpret  $\beta$  in Equation 1 as the causal effect of a same-race match, it must be the case that whether the youth was matched with a same-race mentor is as good as randomly assigned, conditional on the observable characteristics of both the youth and the mentor. Because of the heuristic approach to matching described in Section 2, we include controls for the youth's age, gender, and socioeconomic status and the mentor's age, gender, and educational attainment to account for the decision criteria used by the youth's match specialist in selecting a mentor. To the extent that these controls are correlated with other relevant characteristics, they may also proxy for qualities of the match that we do not observe but that the match specialist does, such as personal interests mentioned during the interview.

One possible threat to identification is that youth are assigned to same-race matches based on some unobserved match characteristics that are correlated with growth in outcomes. Because youth often fill out the baseline survey prior to being matched with a mentor, match specialists may be more likely to allocate same-race mentors to the youth with lower or higher baseline values of certain outcomes.<sup>10</sup> For example, if match specialists expect youth with higher baseline

<sup>9</sup> In the case of multi-racial individuals, we use the first listed race as their primary identity for defining a race-match. For individuals in the "other" category, we determine race congruence using the included subcategories (Asian, Pacific Islander and American Indian) rather than the mentor and youth both being in the "other" category.

<sup>10</sup> In some cases the youth fills out the baseline survey after a potential mentor has been identified but prior to the initial meeting.

scores to benefit more from a same-race mentor our estimates would be positively biased. Table 1 shows the means and standard deviations of youth’s baseline outcomes by match type, the mean difference in scores, the T-statistic of the difference, and the normalized (sample size independent) difference. We find no evidence that youth with higher baseline scores are more likely to be matched with a same-race mentor. Additionally, match specialists may assign youth to same-race matches based on their own demographics or socioeconomic status, or those of the mentor. Table 1 also shows no significance differences in the demographics and socioeconomic status of youth and mentors by match type.

Absent from our match-level specification are agency fixed effects which absorb much of the variation in same-race status, leaving only within-agency variation. In doing so, our results may suffer from time-invariant bias from two sources: region-specific demographic composition and idiosyncratic agency behavior.<sup>11</sup> We account for region-specific factors that might bias our results through the inclusion of other characteristics of the mentor and youth. To the extent that these characteristics correlate with the socioeconomic or demographic characteristics of the local community, these covariates act as a proxy for the unobserved determinants of youth’s growth. In the latter case, some agencies may be more adept at developing successful same-race matches than others by following a set of idiosyncratic best practices in matching or case management. Indeed, local agencies have the freedom to deviate slightly from national policies, possibly to the advantage of youth in same-race matches. Although our discussions with the local agency did not lead us to believe that there were deviations from organization-wide policy that are correlated with the growth among youth, we address this potential bias more systematically with an agency-level model. We describe this strategy in detail in Section 4.3. Finally, we test the robustness of our results to the inclusion of agency fixed effects and find similar effects across both models (see Section 5.1 for a fuller discussion).

It is worth noting that participation in the program is voluntary, so there is likely selection into who enrolls in the program. For that reason, it is important to note that we can only unbiasedly estimate the relative effects of twelve months of mentoring on outcomes for youth in same-race matches relative to those cross-race matches *among youth who elected to enroll in mentoring*. We cannot say what the effect of same-race mentoring is relative to a counterfactual youth who did

---

<sup>11</sup> Agency fixed effects would account for the former as agencies serve the local community within a certain geographic distance.

not receive any mentoring.<sup>12</sup>

#### 4.2. Heterogeneity by Youth Race

We re-estimate Equation 1 interacted with dummies for the race of the youth to determine which youth are most affected by race-matching. The heterogeneous treatment effect model is

$$\begin{aligned} Y_{iat}^F = & \alpha + \beta_B \text{SameRace} \times \text{Black}_{ia} + \beta_W \text{SameRace} \times \text{White}_{ia} \\ & + \beta_H \text{SameRace} \times \text{Hispanic}_{ia} + \beta_O \text{SameRace} \times \text{Other}_{ia} \\ & + \eta Y_{iat}^B + \Gamma' \text{YouthChars}_{iat} + \Theta' \text{MentorChars}_{iat} + \delta_t + \varepsilon_{iat} \end{aligned} \quad (2)$$

where the control set and fixed effects are identical to the average treatment effect model. The only exception is the omission of the mentor's race dummies which are collinear with the youth's race dummies and *SameRace*-youth's-race-dummy interactions. In this model,  $\beta_k$ ,  $k = B, W, H, O$ , is the effect of assigning a same-race mentor relative to a cross-race mentor for a youth of race  $k$ .

#### 4.3. Agency-Level Estimation

To abstract away from any match-level bias, we estimate the agency-level analog of Equation 1 which takes the form

$$\bar{Y}_{at}^F = \alpha + \beta \overline{\text{SameRace}}_{at} + \eta \bar{Y}_{at}^B + \Gamma' \overline{\text{YouthChars}}_{at} + \Theta' \overline{\text{MentorChars}}_{at} + \delta_t + \nu_{at} \quad (3)$$

where  $\bar{Y}_{at}^F$  is the mean of the outcome variable in the follow-up survey at agency  $a$  in year  $t$ .  $\overline{\text{SameRace}}_{at}$  is the fraction of matches in agency  $a$  in year  $t$  that are same-race,  $\bar{Y}_{at}^B$  is the agency-year mean of the outcome variable in the baseline survey, and  $\overline{\text{YouthChars}}_{at}$  and  $\overline{\text{MentorChars}}_{at}$  are agency-year means of the youth and mentor characteristics included in Eq. 1, respectively.

In order to interpret  $\beta$  in Equation 3 as the causal impact of a higher percent of same-race matches, the proportion of observed same-race matches at an agency in any particular year must be as good as randomly assigned, conditional on the included observables. While the agency-level regression model does not suffer from match-level endogeneity concerns, it is susceptible to bias from local economic or social factors such as differences in school spending or the demographic composition which informs the pool of potential mentors and youth. Controlling for the average baseline youth outcome combats bias from social factors, while the proportions of youth receiving

---

<sup>12</sup> For estimates of the impact of youth mentoring enrollment see Grossman and Tierney (1998), Herrera et al. (2011), Park, Liao, and Crosby (2017), Rodriguez-Planas (2012).

free/reduced lunch, youth in a single-parent home, and mentors in each educational attainment bin control for socioeconomic factors. Lastly, we include the proportion of mentors and youth in each race category which controls for local demographic trends over time.

A second, related concern is that the matching heuristic used by agencies is correlated with both the fraction of same-race matches as well as the average outcome at the first follow-up survey. If, for example, agencies with a higher ratio of minority mentors to minority youth more often match promising youth with same-race mentors, our results would be biased. The causal interpretation of Equation 3 rests on the assumption that the fraction of same-race matches is uncorrelated with unobserved agency-specific characteristics.

We similarly estimate Equation 2 at the agency-level by replacing all of the variables with their agency-level means. To estimate the heterogeneous treatment effects at the agency-level, we include the proportion of same-race matches where both mentor and youth are Black, white, Hispanic or in the Other category. We discuss the results of both the estimated average and heterogeneous treatment effects at the match- and agency-levels in the next section.

## 5 — Results

### 5.1. Match-Level Results

Panel A of Table 2 contains the results of estimating Equation 1. The second to last row shows the mean of the dependent variable at baseline among cross-race matches for context. In the match-level regressions, being matched with a mentor of the same race does not appear to improve youth outcomes relative to being matched with a mentor of a different race. The lack of significant impacts can be interpreted in several ways. It may simply be the case that there are no relative benefits to being matched with a mentor of the same race. However, the literature on race-congruency suggests there are positive effects of both same- and cross-race matching (Rhodes et al., 2002). If this is the case, the effects might cancel out, leading to an overall effect close to zero. A final possibility is that certain youth benefit from having a same-race match while others do not. In this case, estimating the effects for the full sample might mask important heterogeneity in the same-race effect.

Panel B contains the results of estimating a model where the same-race indicator variable is interacted with a set of indicators for the youth's race. As the literature suggests, it may be the case that mentoring improves school outcomes for minority youth when paired with a mentor of the

same race. Our results show this to be true in some instances. Black youth in same-race matches saw a 0.0345 point improvement ( $p < 0.05$ ) in their self-perceived school ability relative to those who were mentored by a cross-race mentor. Asian, Pacific Islander and American Indian youth in same-race matches were less truant after a year of mentoring ( $-0.1655, p < 0.001$ ). We find no effects of race-matching on the youth's self-perceived social acceptance for any race category.

Contrary to the existing literature on race-congruency, we find few impacts of race-matching on educational outcomes of youth. Black youth matched to Black mentors had slightly lower grades ( $-0.0034$ ) but this effect is imprecisely estimated. Hispanic youth in same-race matches had slightly lower grades ( $-0.0473, p < 0.05$ ) relative to those in cross-race matches. This could be due to spending extended time with a mentor whose primary language is English. Currie and Thomas (1999) show that the Head Start preschool program had larger impacts on standardized test scores for Hispanic children, and particularly Hispanic children from households where the primary language spoken was Spanish. Unfortunately, we cannot test this directly since we do not observe the youth's language ability or the primary language spoken at home. However, Table 3 shows that the effect on Hispanic youth' grades is driven by their scores in reading, social studies and science, with no effect on math grades. The affected grades are in courses that have a higher marginal return to increased English language ability (e.g. reading comprehension). These results combined with the fact that mentors rarely spend time tutoring their youth provide suggestive evidence of an English language ability mechanism driving the cross-race effects on Hispanic youth' grades. Youth in the Other category had significantly higher math grades ( $0.1795, p < 0.01$ ) as a result of being race-matched, but the effect on overall grades is insignificant. Lastly, we find no effects of race-matching on the youth's expectations for educational attainment.

Turning to the outcomes related to the youth's social interactions, we find a 0.0107 point improvement in the risk attitudes of Black youth matched with a Black mentor relative to those in cross-race matches ( $p < 0.05$ ). We observe a similarly sized effect of 0.0153 for Hispanic youth ( $p < 0.01$ ) and no statistically significant effects for white youth or youth in the Other category. We do not observe any statistically significant effects of race-matching on parental trust. Notably, the small relative decrease in the likelihood that a youth in a same-race match reports having a special adult in their life appears to be driven by Black youth. Because these effects can be interpreted as the relative impact of having a same-race mentor compared to a mentor of another race, these results suggest that Black youth may benefit along some dimensions from having a non-Black mentor.

We test the robustness of these results with two additional specifications. In the first, we include agency fixed effects to account for any unobserved agency-specific and time-invariant omitted variables that are correlated with the outcomes of the youth and the probability they are assigned a same-race mentor. For example, some agencies may be relatively more adept at identifying same-race mentors who will have a greater impact on the youth through a unique screening process. Although all agencies report to the national office, they are allowed to deviate slightly from organization-wide policy to better serve the local community. The drawback of using agency fixed effects is that they may overfit the model by accounting for a common matching heuristic shared by all match specialists at a particular agency. Table A3 shows the results of estimating Equations 1 and 2 with agency fixed effects. The point estimates are similar to those in Table 2 though attenuated and estimated with less precision. While the within-agency estimator is not susceptible to agency-specific bias, it seems to absorb much of the useful identifying variation from estimating the same-race effect across agencies.

In the second specification, we drop any matches where the mentor or youth had two races listed to address measurement error in same-race status. In our sample, there are 1,210 mentors and 3,585 youth who listed two races and thus we cannot be certain of their same-race status. For example, the same-race status between a white, non-Hispanic mentor and a white, Hispanic youth is not obvious. The youth's perception of their own same-race status likely depends on whether they primarily identify as white or Hispanic. Another possibility is that youth may experience an intensity of same-race effects along the race-congruence continuum. In this case, the white, Hispanic youth described above may benefit from having a white, non-Hispanic mentor more than a mentor of a completely different race, but less than having a white, Hispanic mentor. We take a conservative approach to addressing this ambiguity in same-race status by dropping any matches where either participant was multi-racial. Table A4 shows the results of estimating the same-race effect on the subsample of matches where both mentor and youth listed a single race and thus their same-race status is certain. The point estimates are remarkably similar to those in Table 2 but with larger standard errors given the considerable loss of power from dropping around 4,500 observations. From this table we conclude that the effect of race-matching is not driven by multi-racial matches for whom we cannot be certain of their true perception of same-race status.



## 5.2. Agency-Level Results

Table 4 contains results from estimating Equation 3. These results are robust to match-level bias but are susceptible to any contextual effects not captured by our controls for mentor and youth demographics and socioeconomic status. The agency-level results are qualitatively similar in most cases, only scaled by the percent of same-race matches at an agency. This lends credibility to the assumption that there is a negligible amount of selection at the match level. Unless bias from geographic variation is the same direction and magnitude as match specialist selection bias, these results would not persist across specifications. In Panel A, we find increasing the proportion of same-race matching within an agency has no statistically significant impacts on youth' outcomes except in the case of grades. An additional one percent of pairs being race-matched would increase the average grades score by 0.002109 (0.2109/100,  $p < 0.05$ ). When scaled by the average proportion of same-race matches within an agency, the magnitudes of the estimates in Panel A are similar to those in Table 2.

In Panel B, we include the proportion of same-race matches within each race category of the youth. This allows us to examine the impacts of increasing the proportion of same-race matches within youth race subgroups. Although statistically insignificant, we find similarly positive effects on school ability for Black youth. The significant effect of race-matching on grades is driven by white youth in same-race matches: increasing the proportion of white-white matches by one percent leads to a 0.004698 increase in the average grade of white youth ( $p < 0.01$ ). At the agency level, we find that increasing the share of race-matched Hispanic youth increases education expectations by 0.002314 points ( $p < 0.05$ ). Lastly, we find improvements in risk attitudes among Black youth and youth in the Other category, but no effects on parental trust or the prevalence of youth reporting a special adult.

Taken together, Tables 2 and 4 show that race-matching generates modest improvements in the youth's self-perceptions and problem behavior, primarily among Black and Hispanic youth. We find minor improvements in course grades and no impacts on educational expectations, unlike the broader literature on race-congruency. This is perhaps unsurprising given that the program is focused on social mentoring. Mentors often spend time bonding with their youth in leisurely activities rather than tutoring. Indeed, the organization described their program as primarily impacting the social and emotional of development of the youth. The lack of findings for educational outcomes implies that pure role-modeling is not the primary driver of race-

congruence effects on tests scores, high school completion or college enrollment for racial or ethnic minority youth. Outside of the classroom, race-matched mentoring appears to have the highest marginal impacts on the non-cognitive outcomes of youth.

## **6— Conclusion**

In this paper, we estimated the effect of same-race mentorship relative to cross-race mentorship on the outcomes of youth who participated in mentoring for twelve months. We found that youth who were assigned a same-race mentor had almost no improvements relative to those assigned a cross-race mentor, on average. It is possible that both same- and cross-race mentoring have positive impacts for certain youth and negate each other when averaged across the entire sample. Heterogeneity analysis by the race of the youth revealed this is somewhat the case. Same-race matching improved self-perceived school ability for Black youth, truancy for youth in the Other category, and risk attitudes for both Black and Hispanic youth. On the other hand, Hispanic youth in cross-race matches had slightly higher grades after a year of mentoring, and Black youth in cross-race matches were more likely to identify a special adult in their life. It may also be the case that race-matching improves race-relevant outcomes. For example, youth in same-race matches may have better self-perceptions of their race or ethnicity, may have a more positive racial or ethnic identity or may better cope with experienced racial or ethnic discrimination. We cannot conclude that race-matching is not an important determinant for such outcomes as we do not observe them in our data.

Youth mentorship has been shown to have significant positive effects on a range of outcomes for children, and race-congruence is believed to be an important determinant of this success. We contribute to the literature on race-congruence by showing there are potential benefits to both same- and cross-race matching. Furthermore, when full race-matching is not feasible, organizations must choose how to allocate the scarce supply of eligible mentors to youth. We showed that certain groups benefit from race-matching along different dimensions. This heterogeneity in the same-race premium as well as identifying the scenarios when cross-race benefits outweigh same-race benefits are critical for understanding how to efficiently allocate minority mentors in the presence of supply constraints. Our results suggest that policy makers in areas with higher proportions of racial and ethnic minorities, should consider the additional benefits of policies that lessen these supply constraints for youth mentoring programs. For example, the State of

Colorado offers tax credits equal to 50% of donations to the BBBS program.<sup>13,14</sup>

Although we identified for whom same-race mentoring is most impactful, more research is needed to understand the mechanisms behind these effects. Race-congruence seems to impact youth's self-perceptions more than academic performance or attitudes towards adults. This suggests that mentorship improves the youth's self-confidence but may not impact their academic skill or perceptions of authority. This is not unexpected as mentors in community-based matches rarely spend time helping their youth study or complete homework. In addition, we focused on the first twelve months of mentoring. It may be the case that affecting the youth's academic ability or worldview takes more than one year, and further research is needed to estimate the causal impacts on longer term outcomes. Finally, we are not able to study the impact of same-race matching in mentoring relative to no mentoring whatsoever because our data consist only of successful matches. More work is needed to make credible claims about the level effects of race-congruence compared to non-mentored youth.

---

<sup>13</sup> This policy recommendation relies on the assumption that the additional resources would increase the number of minority mentors, either through matches lasting longer (i.e. from greater support) or from having more minority mentors enter the program (i.e. more flexibility in training and potentially recruitment efforts).

<sup>14</sup> <https://www.colorado.gov/pacific/sites/default/files/Income35.pdf>, Accessed April 6th, 2021.

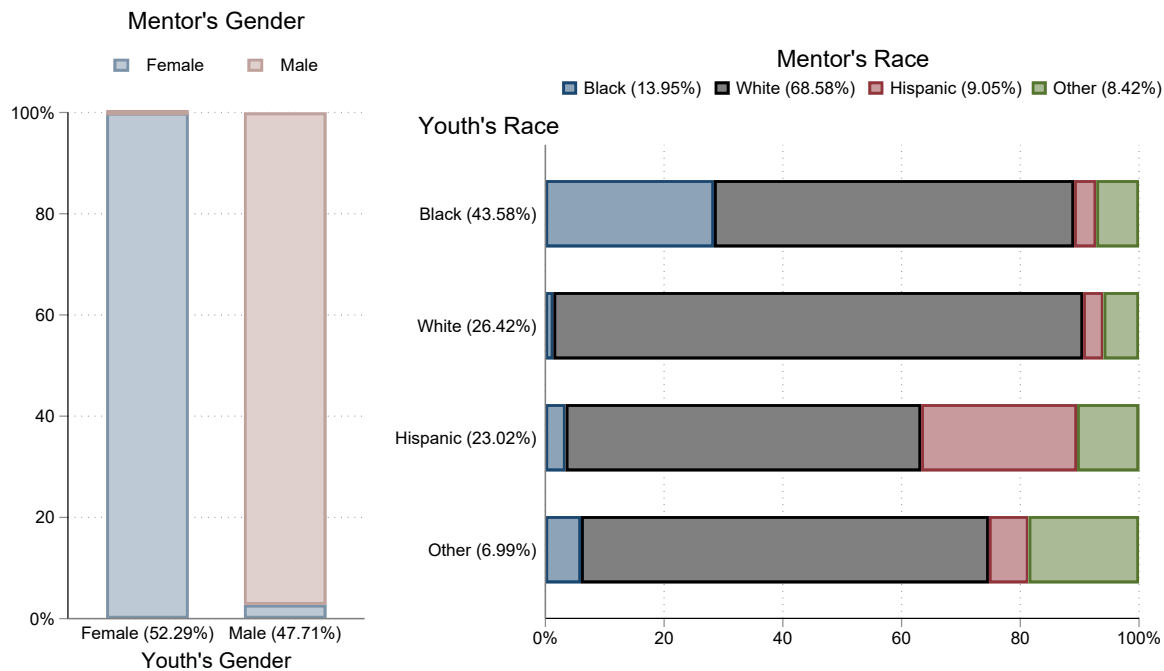
## References

- Armsden, G. C., & Greenberg, M. T. (1987). The inventory of parent and peer attachment: Individual differences and their relationship to psychological well-being in adolescence. *Journal of Youth and Adolescence*, 16(5), 427–454.
- Bates, L. A., & Glick, J. E. (2013). Does it matter if teachers and schools match the student? racial and ethnic disparities in problem behaviors. *Social Science Research*, 42(5), 1180–1190.
- Brown, B. B., Clasen, D. R., & Eicher, S. A. (1986). Perceptions of peer pressure, peer conformity dispositions, and self-reported behavior among adolescents. *Developmental Psychology*, 22(4), 521.
- Bruce, M., & Bridgeland, J. (2014). The mentoring effect: Young people's perspectives on the outcomes and availability of mentoring.
- Currie, J., & Thomas, D. (1999). Does head start help hispanic children? *Journal of Public Economics*, 74(2), 235–262.
- Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, 86(1), 195–210.
- Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, 95(2), 158–165.
- DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? a systematic assessment of the evidence. *Psychological Science in the Public Interest*, 12(2), 57–91.
- Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45, 44–52.
- Falk, A., Kosse, F., & Pinger, P. (2020). Mentoring and schooling decisions: Causal evidence.
- Grossman, J. B., & Tierney, J. P. (1998). Does mentoring work? an impact study of the big brothers big sisters program. *Evaluation Review*, 22(3), 403–426.
- Harbatkin, E. (2021). Does student-teacher race match affect course grades? *Economics of Education Review*, 81, 102081.
- Harter, S. (1982). The perceived competence scale for children. *Child Development*, 87–97.
- Herrera, C., Grossman, J. B., Kauh, T. J., & McMaken, J. (2011). Mentoring in schools: An impact study of big brothers big sisters school-based mentoring. *Child Development*, 82(1), 346–361.

- Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5–86.
- Kofoed, M. S., & McGovney, E. (2019). The effect of same-gender or same-race role models on occupation choice evidence from randomly assigned mentors at west point. *Journal of Human Resources*, 54(2), 430–467.
- Ouazad, A. (2014). Assessed by a teacher like me: Race and teacher assessments. *Education Finance and Policy*, 9(3), 334–372.
- Park, H., Liao, M., & Crosby, S. D. (2017). The impact of big brothers big sisters programs on youth development: An application of the model of homogeneity/diversity relationships. *Children and Youth Services Review*, 82, 60–68.
- Rhodes, J. E., Reddy, R., Grossman, J. B., & Maxine Lee, J. (2002). Volunteer mentoring relationships with minority youth: An analysis of same-versus cross-race matches 1. *Journal of Applied Social Psychology*, 32(10), 2114–2133.
- Rodriguez-Planas, N. (2012). Longer-term impacts of mentoring, educational services, and learning incentives: Evidence from a randomized trial in the united states. *American Economic Journal: Applied Economics*, 4(4), 121–39.
- Van Ewijk, R. (2011). Same work, lower grade? student ethnicity and teachers' subjective assessments. *Economics of Education Review*, 30(5), 1045–1058.

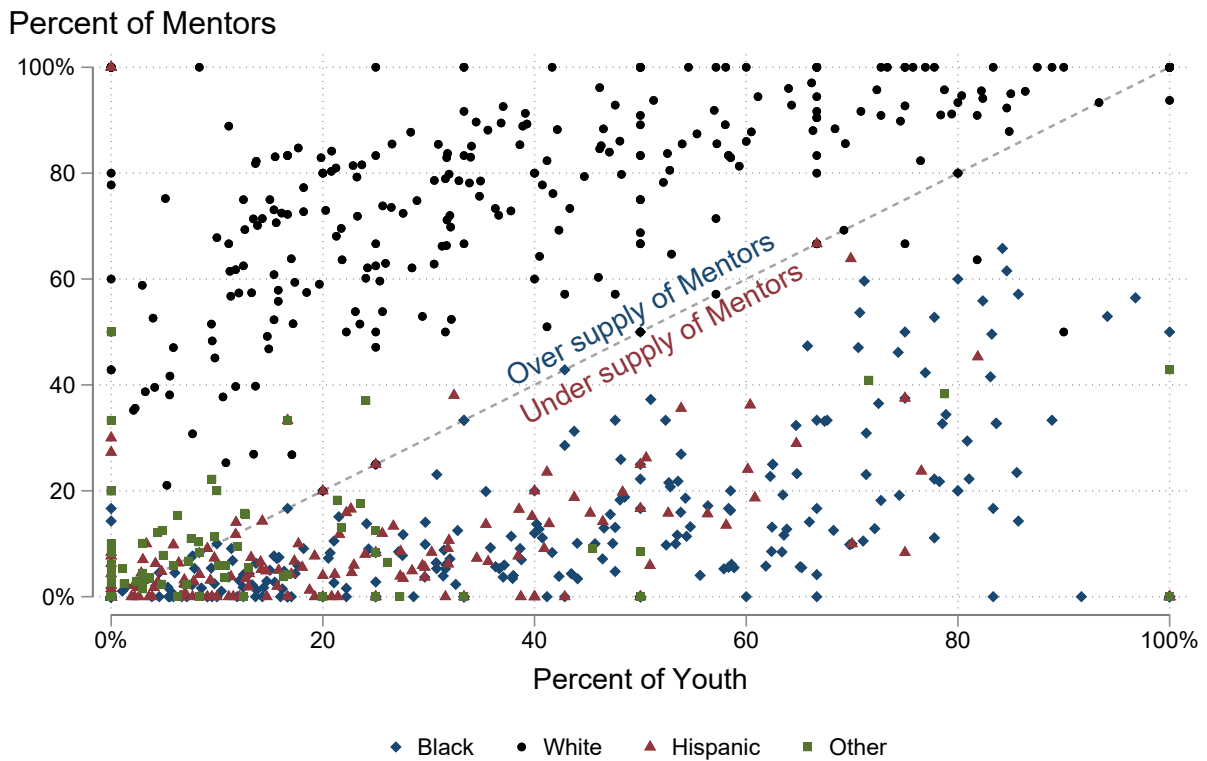
## Tables and Figures

**Figure 1 — Percent of Matches by Mentor/Youth Race and Gender**



*Notes:* Left panel shows the percent of youth by gender matched to female and male mentors. Percent of full sample that are female or male youth shown in parentheses. Right panel shows the percent of youth by race matched to a mentor of a certain race. Percent of the full sample of youth that are of each race are shown in parentheses along the vertical axis. Percent of the full sample of mentors that are of each race are shown in parentheses in the legend. The Other category includes Asian, Pacific Islander, and American Indian. Within the Other-Other cell, same-race is defined using the associated subcategories. The height of each colored portion of the bar shows the proportion of youth matched to a male or female mentor, by the gender of the youth.

**Figure 2 — Variation in Racial Composition of Mentors and Youth by Agency**



*Notes:* Each point represents the proportion of mentors and youth of a particular race within a local agency. The upper diagonal represents all of the agencies where the proportion of mentors in a particular category is greater than the proportion of youth in the same category, and vice versa. The Other category includes Asian, Pacific Islander and American Indian, but not necessarily the relative proportion of the subcategories.

**Table 1 — Balance of Baseline Characteristics by Same-Race Status**

	Same-Race			Cross-Race			Difference		
	Mean	SD	N	Mean	SD	N	Mean	T-stat	Mean/SD
<b>Youth Outcomes</b>									
Social Acceptance	2.84	0.65	12,427	2.91	0.62	16,533	-0.07	-9.90	-0.08
School Ability	2.92	0.60	12,443	2.94	0.59	16,568	-0.01	-2.06	-0.02
Education Expectations	3.55	0.67	12,564	3.62	0.62	16,764	-0.07	-9.60	-0.08
Grades	3.70	0.80	12,414	3.72	0.77	16,510	-0.02	-2.22	-0.02
Risk Attitudes	3.85	0.26	12,507	3.85	0.27	16,692	-0.00	-0.33	-0.00
Parental Trust	3.57	0.58	12,545	3.61	0.55	16,718	-0.04	-6.22	-0.05
Truancy	2.09	0.82	12,527	2.09	0.85	16,668	-0.00	-0.31	-0.00
Special Adult (=1)	0.56	0.50	12,397	0.59	0.49	16,474	-0.03	-4.96	-0.04
<b>Youth Characteristics</b>									
Male (=1)	0.47	0.50	12,653	0.48	0.50	16,882	-0.01	-1.28	-0.01
Age	11.25	1.83	12,649	11.15	1.83	16,870	0.10	4.67	0.04
Free-Reduced Lunch (=1)	0.74	0.44	12,653	0.81	0.39	16,882	-0.07	-14.34	-0.12
Single-Parent HH (=1)	0.67	0.47	12,653	0.70	0.46	16,882	-0.02	-4.39	-0.04
Two-Parent HH (=1)	0.21	0.41	12,653	0.20	0.40	16,882	0.01	2.45	0.02
<b>Mentor Characteristics</b>									
Male (=1)	0.46	0.50	12,653	0.47	0.50	16,882	-0.01	-1.15	-0.01
Age	37.89	12.08	12,652	36.49	11.15	16,881	1.41	10.34	0.09
Less than High School (=1)	0.01	0.09	12,653	0.01	0.07	16,882	0.00	3.17	0.03
High School Graduate (=1)	0.06	0.24	12,653	0.04	0.20	16,882	0.02	8.25	0.07
Some College (=1)	0.22	0.42	12,653	0.17	0.38	16,882	0.05	10.59	0.09
Associate Degree (=1)	0.07	0.25	12,653	0.05	0.21	16,882	0.02	6.53	0.05
Bachelor's Degree (=1)	0.43	0.50	12,653	0.51	0.50	16,882	-0.07	-11.94	-0.10
Advanced Degree (=1)	0.21	0.41	12,653	0.23	0.42	16,882	-0.02	-4.24	-0.04
<b>Match Characteristics</b>									
Match Length (mos)	34.87	19.79	12,653	34.98	19.83	16,882	-0.11	-0.48	-0.00

*Notes:* Means, standard deviations and sample sizes are calculated from the analytical sample of formed matches by same-race status. Same-race status is defined using the specific race recorded for the mentor and youth. In the case of multi-racial individuals, the first listed race is used for matching. All outcomes shown are the baseline values. The last three columns are the difference in means across groups, the T-statistic of the difference and the standardized difference, respectively. The standardized difference, Mean/SD, is the difference in means divided by the standard deviation of the difference (see Imbens and Wooldridge (2009)).



**Table 2 — Same-Race Impacts on Youth's Follow Up Outcomes at the Match Level**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	School Experience			Education		Social Experience		
	Social Acceptance	School Ability	Truancy	Grades	Education Expectations	Risk Attitudes	Parental Trust	Special Adult
<b>Panel A: Average Same Race Effect</b>								
Same Race	-0.0088	-0.0015	-0.0128	0.0145	-0.0181	0.0068	-0.0125	-0.0135
	(0.0120)	(0.0160)	(0.0168)	(0.0170)	(0.0128)	(0.0039)	(0.0104)	(0.0077)
<b>Panel B: Same Race Effect by Race of Youth</b>								
Same Race × Black	-0.0055	0.0345*	-0.0237	-0.0034	0.0112	0.0107*	0.0038	-0.0272***
	(0.0093)	(0.0138)	(0.0195)	(0.0158)	(0.0121)	(0.0053)	(0.0120)	(0.0060)
Same Race × White	-0.0004	-0.0143	0.0264	0.0243	-0.0051	-0.0075	-0.0304	-0.0061
	(0.0211)	(0.0185)	(0.0292)	(0.0292)	(0.0234)	(0.0063)	(0.0162)	(0.0118)
Same Race × Hispanic	-0.0218	-0.0165	-0.0406	-0.0473*	-0.0179	0.0153**	-0.0219	-0.0050
	(0.0146)	(0.0195)	(0.0279)	(0.0240)	(0.0178)	(0.0055)	(0.0142)	(0.0113)
Same Race × Other	-0.0190	0.0375	-0.1655***	0.1183	-0.0150	0.0179	0.0062	-0.0143
	(0.0457)	(0.0807)	(0.0469)	(0.0821)	(0.0557)	(0.0133)	(0.0312)	(0.0276)
Baseline Mean of Cross-Race	2.9136	2.9359	2.0949	3.7190	3.6227	3.8543	3.6100	0.5915
N	28,604	28,633	28,893	28,273	29,114	28,864	28,951	28,574

*Notes:* \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Standard errors in parentheses are clustered at the agency level. All regressions include controls for the youth's baseline outcome, gender, race, free/reduced lunch status, and single-parent home status, mentor's gender and race, as well as fixed effects for the youth's age at follow up, mentor's age and education at follow up, and calendar year. Panel B omits the controls for mentor's race to avoid collinearity with the interaction terms and the include controls for youth's race. Baseline mean of cross-race is the mean of the outcome at baseline among the cross-race group.

**Table 3 — Same-Race Impacts on Youth's Individual Course Grades**

	(1)	(2)	(3)	(4)
	Math	Reading	Social Studies	Science
Same Race × Black	-0.0086 (0.0203)	0.0037 (0.0189)	0.0009 (0.0179)	-0.0130 (0.0234)
Same Race × White	0.0283 (0.0358)	0.0007 (0.0377)	0.0203 (0.0342)	0.0406 (0.0432)
Same Race × Hispanic	-0.0172 (0.0307)	-0.0542* (0.0268)	-0.0714* (0.0286)	-0.0795* (0.0335)
Same Race × Other	0.1795* (0.0901)	0.0696 (0.0925)	0.0624 (0.1006)	0.1420 (0.0817)
Baseline Mean of Cross-Race	3.6380	3.7510	3.6648	3.8120
N	29,183	29,164	28,672	28,827

*Notes:* \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Standard errors in parentheses are clustered at the agency level. All regressions include controls for the youth's baseline outcome, gender, race, free/reduced lunch status, and single-parent home status, mentor's gender, as well as fixed effects for the youth's age at follow up, mentor's age and education at follow up, and calendar year. Baseline mean of cross-race is the mean of the outcome at baseline among the cross-race group.

**Table 4 — Same-Race Impacts on Average Youth Follow Up Outcomes at the Agency Level**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	School Experience			Education		Social Experience		
	Social Acceptance	School Ability	Truancy	Grades	Education Expectations	Risk Attitudes	Parental Trust	Special Adult
<b>Panel A: Average Same Race Effect</b>								
Same Race	0.0078	0.0502	-0.0152	0.2109*	0.0319	0.0316	0.1187	0.0253
	(0.0640)	(0.0825)	(0.0817)	(0.0987)	(0.0646)	(0.0175)	(0.0653)	(0.0393)
<b>Panel B: Same Race Effect Within Race of Youth</b>								
Same Race × Black	-0.0216	0.0740	-0.0050	-0.0030	0.0864	0.0454*	0.0211	-0.0279
	(0.0568)	(0.0636)	(0.0834)	(0.0749)	(0.0641)	(0.0182)	(0.0536)	(0.0399)
Same Race × White	0.0573	0.0138	0.0129	0.4698**	-0.1116	0.0335	0.1600	0.1059
	(0.1196)	(0.1549)	(0.1331)	(0.1581)	(0.0946)	(0.0317)	(0.1071)	(0.0747)
Same Race × Hispanic	0.0582	0.1275	-0.1641	0.0393	0.2314*	0.0396	0.1101	0.0399
	(0.0979)	(0.1161)	(0.1363)	(0.1471)	(0.0929)	(0.0312)	(0.1018)	(0.0585)
Same Race × Other	-0.1316	0.2483	0.0407	0.3980	0.4794	0.1357*	0.2561	-0.0872
	(0.2230)	(0.2661)	(0.2570)	(0.2993)	(0.2506)	(0.0611)	(0.2057)	(0.1240)
Baseline Mean of Outcome	2.8227	2.9119	2.0845	3.6496	3.5174	3.8568	3.5969	0.5655
Fraction of Same-Race Matches	0.5069	0.5066	0.5066	0.5020	0.5046	0.5046	0.5043	0.5055
N	1740	1743	1745	1735	1746	1742	1745	1739

*Notes:* \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Robust standard errors in parentheses. All regressions include controls for the average baseline score of all youth at the agency; the within-agency fraction of youth that are: male, in each race category, in each age year bin, on free/reduced lunch, and live in a single-parent home; the within-agency fraction of mentor's that are: male, in each race category, of each age, in each educational attainment bin; and calendar year fixed effects. Panel B omits the controls for the fraction of mentors in each race bin to avoid collinearity with the interaction terms and the fraction of youth in each race bin. Baseline mean is the mean of the outcome at baseline. Fraction of same-race matches is the average proportion of matches within an agency that are same-race.

## Appendix

**Table A1 — Average Baseline and Follow-up Youth Outcomes by Same-Race Status**

	Same-Race		Cross-Race	
	Baseline	Follow-up	Baseline	Follow-up
Social Acceptance	2.839	2.987	2.914	3.079
School Ability	2.922	3.063	2.936	3.082
Truancy	2.091	2.104	2.094	2.114
Education Expectations	3.549	3.607	3.622	3.669
Grades	3.696	3.684	3.716	3.688
Risk Attitudes	3.853	3.899	3.854	3.903
Parental Trust	3.566	3.672	3.608	3.706
Special Adult (=1)	0.563	0.883	0.592	0.888

*Notes:* Shown are unadjusted means of each outcome at pre-match baseline and first follow-up by same-race status.

**Table A2 — Balance Between Matches With and Without Follow-up Surveys**

	Has Follow-up			No Follow-up			Difference		
	Mean	SD	N	Mean	SD	N	Mean	T-stat	Mean/SD
<b>Youth Outcomes</b>									
Social Acceptance	2.88	0.63	28,959	2.89	0.65	43,955	-0.01	-1.71	-0.01
School Ability	2.93	0.59	29,010	2.90	0.60	43,976	0.03	5.60	0.03
Truancy	2.09	0.84	29,194	2.15	0.85	44,155	-0.05	-8.34	-0.04
Grades	3.71	0.78	28,923	3.65	0.80	43,914	0.06	9.86	0.05
Education Expectations	3.59	0.64	29,327	3.55	0.67	44,488	0.04	7.40	0.04
Risk Attitudes	3.85	0.26	29,198	3.83	0.30	44,228	0.02	10.83	0.06
Parental Trust	3.59	0.57	29,262	3.55	0.61	44,341	0.04	9.84	0.05
Special Adult (=1)	0.58	0.49	28,870	0.58	0.49	43,773	0.00	0.96	0.01
<b>Youth Characteristics</b>									
Male (=1)	0.48	0.50	30,466	0.45	0.50	44,993	0.03	6.69	0.04
Age	11.20	1.83	29,518	11.46	1.95	44,781	-0.26	-18.17	-0.10
Free-Reduced Lunch (=1)	0.78	0.42	30,466	0.78	0.42	44,993	0.00	0.23	0.00
Single-Parent HH (=1)	0.69	0.46	30,466	0.69	0.46	44,993	-0.01	-1.38	-0.01
Two-Parent HH (=1)	0.20	0.40	30,466	0.19	0.39	44,993	0.01	3.47	0.02
<b>Mentor Characteristics</b>									
Male (=1)	0.47	0.50	30,466	0.44	0.50	44,993	0.03	7.42	0.04
Age	37.09	11.58	29,532	36.34	11.60	44,828	0.75	8.64	0.05
Less than High School (=1)	0.01	0.08	30,466	0.01	0.09	44,993	-0.00	-2.25	-0.01
High School Graduate (=1)	0.05	0.22	30,466	0.07	0.25	44,993	-0.02	-9.49	-0.05
Some College (=1)	0.19	0.40	30,466	0.23	0.42	44,993	-0.04	-13.42	-0.07
Associate Degree (=1)	0.06	0.23	30,466	0.06	0.24	44,993	-0.01	-2.85	-0.01
Bachelor's Degree (=1)	0.48	0.50	30,466	0.44	0.50	44,993	0.04	10.78	0.06
Advanced Degree (=1)	0.22	0.41	30,466	0.20	0.40	44,993	0.02	8.08	0.04
<b>Match Characteristics</b>									
Match Length (mos)	35.01	19.91	30,466	15.65	14.99	44,993	19.36	152.15	0.80

*Notes:* Means, standard deviations and sample sizes are calculated from the sample of formed matches by follow-up status. All outcomes shown are the baseline values. The last three columns are the difference in means across groups, the T-statistic of the difference and the standardized difference, respectively. The standardized difference, Mean/SD, is the difference in means divided by the standard deviation of the difference (see Imbens and Wooldridge (2009)).

**Table A3 — Match Level Effects with Agency Fixed Effects**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	School Experience			Education		Social Experience		
	Social Acceptance	School Ability	Truancy	Grades	Education Expectations	Risk Attitudes	Parental Trust	Special Adult
<b>Panel A: Average Same Race Effect</b>								
Same Race	-0.0090 (0.0072)	-0.0042 (0.0093)	-0.0013 (0.0116)	-0.0108 (0.0112)	-0.0155 (0.0082)	0.0019 (0.0026)	-0.0112 (0.0075)	-0.0193*** (0.0054)
<b>Panel B: Same Race Effect by Race of Youth</b>								
Same Race × Black	0.0011 (0.0082)	0.0125 (0.0106)	0.0187 (0.0165)	-0.0081 (0.0178)	-0.0078 (0.0120)	0.0005 (0.0033)	0.0004 (0.0102)	-0.0306*** (0.0062)
Same Race × White	-0.0003 (0.0213)	-0.0043 (0.0176)	0.0186 (0.0271)	0.0082 (0.0304)	0.0042 (0.0241)	-0.0032 (0.0063)	-0.0297 (0.0166)	-0.0046 (0.0122)
Same Race × Hispanic	-0.0268 (0.0168)	-0.0311 (0.0186)	-0.0296 (0.0185)	-0.0387 (0.0212)	-0.0329* (0.0167)	0.0069 (0.0050)	-0.0245 (0.0138)	-0.0053 (0.0111)
Same Race × Other	-0.0477 (0.0350)	-0.0308 (0.0342)	-0.1284** (0.0474)	0.0548 (0.0480)	-0.0698 (0.0371)	0.0056 (0.0108)	0.0011 (0.0256)	-0.0266 (0.0211)
Baseline Mean of Cross-Race	2.9138	2.9360	2.0948	3.7190	3.6228	3.8543	3.6102	0.5916
N	28,601	28,630	28,890	28,270	29,111	28,861	28,948	28,571

Notes: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Standard errors in parentheses are clustered at the agency level. All regressions include the controls listed in Table 2 as well as agency fixed effects. Baseline mean of cross-race is the mean of the outcome at baseline among the cross-race group.

**Table A4 — Match Level Effects Dropping Multi-racial Individuals**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	School Experience			Education		Social Experience		
	Social Acceptance	School Ability	Truancy	Grades	Education Expectations	Risk Attitudes	Parental Trust	Special Adult
<b>Panel A: Average Same Race Effect</b>								
Same Race	-0.0133 (0.0093)	0.0112 (0.0121)	-0.0118 (0.0150)	-0.0130 (0.0135)	-0.0006 (0.0096)	0.0075 (0.0040)	-0.0091 (0.0087)	-0.0132* (0.0059)
<b>Panel B: Same Race Effect by Race of Youth</b>								
Same Race × Black	-0.0060 (0.0113)	0.0308* (0.0148)	-0.0025 (0.0199)	-0.0135 (0.0173)	0.0019 (0.0121)	0.0094 (0.0057)	0.0016 (0.0127)	-0.0246*** (0.0064)
Same Race × White	-0.0174 (0.0230)	-0.0309 (0.0219)	0.0272 (0.0344)	0.0156 (0.0324)	0.0107 (0.0245)	-0.0103 (0.0074)	-0.0333* (0.0163)	-0.0034 (0.0134)
Same Race × Hispanic	-0.0259 (0.0170)	-0.0109 (0.0214)	-0.0300 (0.0320)	-0.0534* (0.0271)	-0.0182 (0.0185)	0.0132 (0.0070)	-0.0179 (0.0131)	0.0020 (0.0119)
Same Race × Other	-0.0122 (0.0499)	0.0503 (0.0954)	-0.1464** (0.0504)	0.1380 (0.0961)	0.0378 (0.0619)	0.0110 (0.0166)	-0.0026 (0.0393)	-0.0018 (0.0287)
Baseline Mean of Cross-Race	2.9148	2.9420	2.0934	3.7225	3.6350	3.8562	3.6117	0.5949
N	24,155	24,141	24,383	23,833	24,553	24,359	24,436	24,101

Notes: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Standard errors in parentheses are clustered at the agency level. All regressions include the controls listed in Table 2. Baseline mean of cross-race is the mean of the outcome at baseline among the cross-race group.