NBER WORKING PAPER SERIES

# HOW INFORMAL MENTORING BY TEACHERS, COUNSELORS, AND COACHES SUPPORTS STUDENTS' LONG-RUN ACADEMIC SUCCESS

Matthew A. Kraft
Alexander J. Bolves
Noelle M. Hurd

Working Paper 31257
http://www.nber.org/papers/w31257

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2023

Corresponding Author: Matthew A. Kraft. We are grateful for the helpful comments on our manuscript from Sarah Cohodes, Scott Imberman, Kirabo Jackson, Lars Lefgren, and seminar participants at University College London, Syracuse University, and the Student Experience Research Network. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

How Informal Mentoring by Teachers, Counselors, and Coaches Supports Students' Long-Run Academic Success
Matthew A. Kraft, Alexander J. Bolves, and Noelle M. Hurd
NBER Working Paper No. 31257
May 2023
JEL No. I21,I24,I26,J24

**ABSTRACT**

We document a largely unrecognized pathway through which schools promote human capital development – by fostering informal mentoring relationships between students and teachers, counselors, and coaches. Using longitudinal data from a nationally representative sample of adolescents, we explore the nature and consequences of natural mentoring relationships by leveraging within-student variation in the timing of mentorship formation as well as differences in exposure among pairs of twins, best friends, and romantic partners. Results across difference-in-differences and pair fixed-effect specifications show consistent and meaningful positive effects on student attainment, with a conservative estimate of a 9.4 percentage point increase in college attendance. Effects are largest for students of lower socioeconomic status and robust to controls for individual characteristics and bounding exercises for selection on unobservables. Smaller class sizes and a school culture where students have a strong sense of belonging are important school-level predictors of having a K-12 natural mentor.

Matthew A. Kraft
Brown University
PO Box 1938
Providence, RI 02478
and NBER
mkraft@brown.edu

Alexander J. Bolves
Harvard University
alexander_bolves@g.harvard.edu

Noelle M. Hurd
University of Virginia
nh3v@virginia.edu

Formal education is the principal investment societies make in the human capital of their youth. Decades of evidence now documents the large returns to additional years of schooling (Gunderson and Oreopolous, 2020). However, we still have a limited understanding of why schools serve as engines of human capital development. As Oreopoulos and Salvanes (2011, p.159) explain, much of the economics literature treats schools as a black box where "individuals enter, something happens, and productivity increases." More recent studies have begun to look inside the black box of schooling, focusing on the role of instructional inputs in the education production process such as teachers, curricular materials, and remedial classes (e.g. Chetty et al., 2014; Jackson, 2018; Cortes et al., 2015).

In this paper, we propose and explore a new lens for understanding how formal education promotes human capital formation – one that views schools as incubators of natural mentoring relationships. This framework bridges two largely distinct research traditions in labor economics and developmental psychology to examine the role of relationships in the education production process. The psychology literature defines natural mentorships as caring relationships between nonparental adults and youth that arise out of existing social networks (Rhodes et al., 1992; Zimmerman et al., 2005). Mentors step outside of the boundaries of their primary roles to develop a unique and sustained relationship with individual youth. Studies find that approximately 70-80% of adolescents can identify at least one natural mentor in their life (Beam et al., 2002; Dubois and Silverthorn, 2005a; Hurd et al., 2016; Hurd and Zimmerman, 2014).

Natural mentors can fill a diverse range of roles in students' lives from role models and caring adults to advisors and advocates. Research suggests that natural mentoring relationships can benefit youth through cognitive, social-emotional, and identity development (Miranda-Chan et al., 2016; Rhodes, 2005; Rhodes and Dubois, 2006). Mentors may broaden students' cognitive frameworks by exposing them to new ways of thinking and alternative perspectives. Positive relationships with mentors can help develop social-emotional skills by modeling effective communication and providing a sounding board to help youth better regulate their emotions (Deutsch et al., 2020; Hurd and Sellers, 2013; Sánchez et al., 2008; Van Dam et al., 2018). Mentors can expand adolescents' self-perceptions and aspirations of who they might become by exposing them to a greater range of "possible selves" (Hurd, et al., 2012; Rhodes and Dubois, 2006).

There is good reason to think that schools are primary sites where informal mentoring relationships develop. Outside of family members, school personnel such as teachers, counselors, and coaches often have the most regular contact with youths' day-to-day lives. These frequent interactions forge bonds that can lead to natural mentoring relationships. Moreover, school-based mentors are uniquely positioned to help students overcome obstacles in school and guide them towards higher education. School personnel may also play an important role in expanding the social capital of underserved youth by increasing access to job opportunities and exposing them to broader social networks that they might not otherwise have access to within their familial, neighborhood, and social circles (Granovetter, 1973).

We leverage longitudinal data from a large, nationally representative sample of adolescents to document the frequency, nature, and school-level correlates of school-based mentoring relationships and to explore their consequences for students' human capital formation. Understanding the distribution and effects of these mentoring relationships across students and schools has important implications for educational equity and opportunity. While natural mentoring relationships could be equally advantageous to all students, they may not be equally accessible to students attending, for example, under-resourced or over-crowded schools. Natural mentors might also serve as a compensatory resource, yielding greater benefits to students facing economic or structural disadvantages, or as a complementary one, adding to the privilege and opportunities enjoyed by more advantaged students (Erickson et al., 2009).

Experimental evidence from formal youth mentoring programs documents meaningful effects on academic performance, educational attainment, and social-emotional skills for students from socioeconomically disadvantaged backgrounds (Grossman and Tierney, 1998; Herrera et al, 2011; Kosse et al., 2020; Resnjanskij et al., 2021). However, identifying the causal effect of informal mentorships poses a greater empirical challenge. Unlike studies of formal mentorship programs such as Big Brothers/Big Sisters, randomized field trials are infeasible for natural mentorships because – by definition – they occur organically. The existing literature on natural mentoring is limited to correlational studies and those that make strong conditional independence assumptions based on observable characteristics (Van Dam et al., 2018). Such estimates are subject to potential biases given the non-random selection processes through which mentees and mentors mutually reciprocate these relationships (Gowdy et al., 2020).

We approach this challenge by leveraging detailed transcript data and within-student variation to estimate effects on short-run outcomes in high school. Our difference-in-differences two-way fixed effects (TWFE) model removes all fixed individual student characteristics, identifying causal effects under the assumptions of common trends and that natural mentorship formation is not confounded with other concurrent shocks to students' academic performance. We then estimate effects on short-run and long-run education outcomes by comparing differences in exposure among twins who attended the same high school. Twins fixed effect (FE) models purge unmeasured family heterogeneity and are a widely-used approach for evaluating the effects of education, family background, individual characteristics, and early childhood experiences on human capital development when randomization is infeasible (e.g. Head Start: Deming, 2009; school quality: Autor et al, 2016; socioeconomic status: Autor et al, 2019; ADHD: Currie, 2006; birth weight: Black, Devereux, & Salvanes al., 2007; Figlio et al., 2014; child maltreatment: Currie & Tekin, 2012). This strategy identifies the effect of school-based natural mentors under the condition that there are no unmeasured individual or personality characteristics within twin pairs that are correlated with both mentorship formation and human capital development. We offer a range of evidence of the plausibility of the assumptions underpinning both our difference-in-differences TWFE and twins FE models.

Natural mentorships between students and school personnel are relatively common. We find that over 15% of adolescents identify a K-12 teacher, counselor, or coach as their most important mentor. Our descriptive analyses illustrate how the frequency of school-based natural mentorships varies meaningfully across students and schools. Similar to prior research, we find that adolescents who are Black, Latinx, and from lower socioeconomic (SES) backgrounds are less likely to report having a school-based natural mentor. We also document for the first time that the prevalence of school-based mentorships ranges considerably across high schools, with mentoring rates more than twice as high in some schools compared to others. Features of the school environment such as smaller class sizes and a culture where students have a strong sense of belonging are important predictors of this variation across schools.

We find a consistent pattern of significant positive effects across both our TWFE and twins FE estimators as well as in models that estimate effects using best-friend, romantic partner, and school FE. In the short run, having a school-based mentor lowers rates of course failure in high school, while increasing credits earned and GPA. In the long run, we find that students who benefit from school-based natural mentors are substantially more likely to attend college and complete roughly two-thirds more of a year of formal education. Our findings are also consistent with previous research suggesting that natural mentoring relationships can play a compensatory role for youth facing disadvantages due to structural or economic factors (Erickson et al., 2009). Heterogeneity analyses suggest that school-based natural mentors benefit students of all backgrounds, but are most beneficial for students from lower-SES backgrounds. While these results are not always precisely estimated, the general pattern is consistent across short- and long-run education outcomes. We find limited evidence of heterogeneity across other aspects of student identity, such as race/ethnicity, gender, and their intersections.

The consistency of our results across models that draw on very different samples, identifying variation, and assumptions serves as a first-order robustness check. We then demonstrate that there is no evidence of systematic bias in exposure to school-based natural mentors within twin, best-friend, or romantic partner pairs along a range of individual and personality characteristics. We also show that controlling for individual and personality characteristics that are correlated with mentorship formation leaves our results fundamentally unchanged. We conduct an additional bounding exercise following Oster (2019) and find that our estimates remain economically meaningful even after allowing for a high degree of selection on unobservables. For example, we estimate a 9.4 percentage point increase in the probability of attending college assuming selection on unobservables is of equal magnitude as selection on observables. Together, our analyses provide the most credible empirical evidence to date on the effect of school-based natural mentors on adolescents' human capital development.

Although isolating exact mechanisms is challenging, several pieces of evidence suggest that the effects we find are a product of mentoring relationships that extend well beyond the classroom rather than the direct effects of teachers' academic instruction. These school-based relationships are long-lasting, with students reporting that their mentors played important roles in their lives for more than five years on average. Two-thirds of students reported that mentors helped them with life development skills such as finding direction in life, setting the right priorities, navigating life crises, and making good decisions. The magnitude of our estimates also remain largely consistent when we isolate the effects of having a sports coach as a mentor.

We build on and contribute to several literatures with this work. First, our paper conceptualizes and develops original evidence in support of a largely unrecognized mechanism through which schools promote human capital development. Second, we provide the most credible evidence to date on the effect of natural mentorships on adolescents' human capital formation. Third, we focus specifically on school-based natural mentorships, while the natural mentorship literature largely focuses on mentors as a general group with a few important exceptions (DuBois and Silverthorn, 2005b; Erickson et al., 2009; Fruiht and Wray-Lake, 2013). This distinction is important because prior evidence suggests that selection mechanisms and affected outcomes differ substantially across natural mentor types (e.g., familial vs. non-familial; Raposa et al., 2018).

More broadly, we contribute to the teacher and counselor effects literatures by illustrating an important pathway through which school personnel may affect students' outcomes outside of their traditional roles and job responsibilities (Chetty et al., 2014; Mulhern, 2020). Our estimates of the effects of natural mentors on students' longer-term educational attainment are substantially larger than the effects found in the teacher effectiveness literature (Chetty et al., 2014; Petek and Pope, 2023). The magnitude and likely mechanisms underlying the effects we find are most closely related to the literature on same-race teachers (Dee, 2004). Our estimates are comparable to the effect of same-race teachers on college enrollment for Black students, and we similarly find that students from economically disadvantaged backgrounds benefit most (Gershenson et al., 2022). Research suggests that same-race teacher effects operate through effective instruction such as culturally relevant pedagogies and role model effects where Black teachers serve as academic mentors who can reveal the hidden curriculum of pursuing postsecondary education (Blazar, 2021).

We organize the remainder of the paper by first describing the data we use. We then provide detailed descriptive statistics on which students report having a K-12 natural mentor and the nature of their mentoring relationships. These descriptives serve to highlight the non-random selection into K-12 mentoring relationships. We then discuss our econometric approach motivated by this selection challenge and present our primary results and robustness checks. Finally, we conduct a range of exploratory analyses which point to specific ways in which policymakers and administrators might foster the formation of natural mentoring relationships in schools.

## I. Data

Our analysis draws on the National Longitudinal Study of Adolescent to Adult Health (Add Health), a study that began following a nationally representative sample of middle and high schoolers (ages 12-19) in the 1994-95 academic year. Over the ensuing three decades, these individuals participated in five waves of intensive in-home interviews as they transitioned into adulthood (Harris et al., 2013). Critical for our focus on school-based mentors, Add Health implemented a stratified sampling design, selecting 105 schools from a stratified list of more than 26,000 high schools across the nation.[1] Schools were then further stratified by grade level

[1] Candidate schools were stratified by region, urbanicity, school type (public, private, parochial), ethnic mix, and size. The probability of being selected was proportional to a school's enrollment.

and gender, and a random sample of about 20 students was taken from each stratum resulting in a final sample of over 20,000 students in grades 7-12 in the 1994-95 school year (Chen and Harris, 2020). The response rate exceeded 80% across all five survey waves, providing a consistent analytic sample across the 25-year span.

We report summary statistics for the weighted nationally representative core sample in Table 1 column 1. The majority of students are white (65%), with Black students comprising 15%, Latinx students comprising 12%, and Asian students comprising 3% of the sample. More than 90% of students were born in the U.S. and approximately one-third have at least one parent who completed college. Participants tended to reside in neighborhoods (census tracts) that were mostly white with roughly a quarter of residents not having earned a high school diploma by age 25 and another quarter having earned a college degree by that same age.

*Natural Mentors.* The third wave of data (when respondents were 18-26 years old) included a question about natural mentoring relationships during adolescence. Specifically, participants were asked "*other than your parents or step-parents, has an adult made an important positive difference in your life at any time since you were 14 years old?"* Respondents were only allowed to identify one individual and were directed to recall the *most* impactful non-parental adult in their lives. Respondents who identified a natural mentor answered a series of follow-up questions aimed at characterizing the relationship. These items captured information about a mentor's gender, how the two met, how long the relationship lasted, the level of closeness in the relationship, how frequently the two interact, and other features.[2]

Add Health also asked how respondents knew their natural mentors prior to forming the mentorship. We categorize mentors into one of three categories: 1) teachers/guidance counselors

[2] Respondents provide the integer age when they first met their mentor which we use to identify the grade students meet their mentors. We assume students and mentors meet in the earliest academic year for a given age.

and coaches/athletic directors that students met before their expected high school graduation date (to distinguish K-12 mentors from those developed in higher education), 2) non-school-based mentors met before expected on-time high school graduation, and 3) mentors met after high school. Our analyses focus exclusively on the first category of K-12 school-based mentors, but we include indicators for the other two natural mentor categories as controls in models comparing outcomes across students.

Another key survey question asked respondents to describe in an open-ended manner "What did [your natural mentor] do to help you?" Add Health coded responses as describing *behaviors* of mentors and *domains* of mentoring.[3] There are eight categories of mentor behaviors which describe the specific interactions between mentors and mentees: giving guidance and advice, providing emotional nurturance, giving practical/tangible help, providing a parental figure, providing a friend figure, providing a role model figure, spending time together, and other responses which do not fit into the primary categories. The seven domains describe the broad areas of students' lives that were influenced by the natural mentorship: personal development, family and household, religion, finances, employment, education, and quality-time in leisure and/or sports, and other responses which mention a domain that is not characterized by these primary categories.

*Twin Subsample.* Add Health oversampled identical and fraternal twins by always including both siblings. The full Add Health sample of twins includes 1,565 students (one set of triplets), which we restrict to an analytic sample of 1,213 students with valid outcomes and demographics. The response rate for twins exceeds 90% across all five waves of interviews (Chen and Harris, 2020; Harris et al., 2013). In Appendix Table B1, we provide summary

[3] Add Health developed a coding scheme where responses were coded into all applicable behaviors and domains. We provide more information on the coding process including inter-rater reliability measures in Appendix A.

statistics for the core sample and twins subsample, and in columns 1-3 we compare twins to non-twins. These two groups are similar with small differences in racial representation. Compared to non-twins, twins are more likely to be Black and Latinx and less likely to be white.

*Friend Networks.* Add Health Waves I and II asked students to nominate their closest friends from each gender, one male best friend and one female. If nominated friends went to a school in the Add Health sample, then friends were added to the core sample.[4] Of the nearly 62,000 friendship nominations from Waves I and II, we drop just under one-third because the nominated friend goes to a school outside of the Add Health sample. Next, we create requited best-friend pairs by identifying friends who mutually nominated one another. Our final best friends analytic sample includes 1,378 students. In Appendix Table B1 columns 4-6, we describe the sample of requited best friends compared to students not in requited best friendships. These samples are similar to one another along observable characteristics with the exception of more Asian students and fewer white students in requited best friendships.

*Romantic Partner Networks.* Add Health Waves I and II also asked students to identify up to three individuals they had a romantic relationship with in the previous 18 months. Respondents were not asked to order romantic partnership nominations in any systematic way (e.g., recency, closeness, duration, etc.). Partners who attended Add Health sample schools were added to the core sample. We conduct an iterative matching process for assigning students to a single requited romantic partnership.[5] This iterative process allows students to be associated with, at most, one romantic partnership. Our final analytic sample of romantic partners contains

---

[4] Roughly one-third of respondents were asked to name up to 5 friends from each gender. Respondents were asked to rank these nominations such that the first nomination is one's closest friend. Following Duncan et al., 2001., we only use the top nominated friend from each gender for individuals who nominated five ranked friends.
[5] Our iterative process is a modified Gale-Shapley algorithm, described further in Appendix C.

548 individuals. Much like best-friend pairs, Appendix Table B1 columns 7-9 show that students with romantic partners are quite similar to the remaining student sample.

*Socioeconomic Status.* We create a broad measure of student socioeconomic status using Wave I data on individual and neighborhood characteristics. Individual characteristics include household income, highest education of either parent, whether either parent is employed full-time, and whether students were covered by health insurance. Census tract measures include: 1) average household income, 2) unemployment rate, and the proportions of households that are 3) receiving welfare assistance, 4) without a high school diploma by age 25, 5) with a college degree by age 25, and 6) owner occupied. We construct our SES composite measure by reversing scales with negative valences, standardizing each measure to have a mean of one and unit variance, and taking individual student averages across these measures. Finally, we standardize this average across the full Add Health sample to arrive at our SES measure.

*Outcomes.* Our primary outcomes of interest are students' educational achievement and attainment. We use detailed data from student transcripts to measure several academic outcomes in each year of high school. These include annual GPA, course failure rate, and the number of length-adjusted year-long courses passed. We also examine educational attainment by Wave IV (ages 24 to 32) using indicators for whether a student attended college, attended a college with a selective admission process, and a measure of total years of education.[6,7] Appendix Table B2 provides nationally representative summary statistics and descriptions for all our outcomes.

---

[6] Our measure of college selectivity is based on Barron's selectivity index.
[7] We identify years of educational attainment based on respondents' highest reported level of attained education. We code the highest attained levels as follows: "8th grade or less" we code as 8 years; "some high school" as 11 years; "high school graduate" as 13 years; "some vocational/technical training (after high school) as 13.5 years; "completed vocational/technical training (after high school)" as 14; "some college" as 15; "completed college (bachelor's degree)" as 17 years; "some graduate school" and "some post baccalaureate professional education (e.g., law school, med school, nurse)" as 18 years; "completed a master's degree" as 19 years; "some graduate training beyond a master's degree" and "completed post baccalaureate professional education (e.g., law school, med school, nurse)" as 20 years; and "completed a doctoral degree" as 22 years.

*Data Limitations.* The timing and nature of the questions asked by Add Health about students' natural mentors present two key limitations to our analyses. First, Add Health's decision to restrict respondents to naming only the single-most influential non-parental adult in their lives shapes the interpretation and external validity of our findings. Recent studies show that students who can identify one natural mentor in their lives can usually identify two or more (Gowdy et al., 2020; Hurd, et al., 2018). Thus, our descriptive data provide a lower-bound estimate of the share of students who develop mentoring relationships with K-12 school personnel. Our model-based estimates are also best interpreted as the relationships we might expect to find for the most impactful school-based natural mentors.

Second, the retrospective nature of the question identifying natural mentors asked in Wave III, when respondents are ages 18 to 26, presents the possibility of recall bias. This could bias our analyses if individuals' likelihood of identifying a natural mentor is systematically influenced by their experienced life outcomes. We provide evidence in our robustness section that this potential threat is unlikely to be a first-order concern in our context.

## II. Describing School-Based Natural Mentoring Relationships

The formation of natural mentoring relationships is a voluntary and informal two-way matching process. Students can seek out and be receptive to mentoring, and school personnel can offer and respond to invitations to be a mentor. Here we describe how often these relationships occur, who forms these relationships, and what type of things mentors do to support students. These descriptive findings elucidate the relevant selection pathways which inform our identification approach described in the methods section below.

### II.a. How Common Are School-Based Mentorships?

We find that school-based natural mentorships are relatively common among youth in our nationally representative sample, with 15.2% of respondents reporting their most impactful natural mentor was a teacher, counselor, or coach. These school-based mentors compose a quarter of all reported natural mentorships and are the second largest source of natural mentors behind only family members which comprise 34% of all mentorships. About 90% of school-based natural mentors are teachers or guidance counselors, and the remaining are coaches or athletic directors. Students were most likely to meet the school-based mentors they identified towards the end of $9^{th}$ or beginning of $10^{th}$ grade (see Appendix Figure B3).

**II.b. Which Students Engage in School-Based Mentorships?**

Previous research suggests that adults may be more likely to form mentoring relationships with young people when they share common demographics (e.g., race/ethnicity, gender) or life experiences (Ensher and Murphy, 1997; Stanton-Salazar and Dornbusch, 1995). Given that the U.S. high school teacher workforce is overwhelmingly white (79%), female (59%), and from middle and upper-middle class SES backgrounds (Hussar et al., 2020; Jacinto and Gershenson, 2021), we might expect students of color, students from economically disadvantaged backgrounds, and male students to be less likely to develop informal mentorships with school personnel.

We do find that white students and Asian male students are substantially more likely to report having a K-12 natural mentor than their Black and Latinx peers. As shown in Figure 1, Black and Latina female students each had rates of K-12 mentorship of about 10%, while Black and Latino males reported slightly higher rates of about 12%. Comparatively, white students reported having school-based mentors about 15% of the time, and more than 20% of Asian male students identified a K-12 natural mentor.

When we compare the socioeconomic status of students with school-based mentors to those with no mentors, we also find clear patterns of privilege-based selection. In Table 1 columns 2-4, we provide descriptive statistics for students who identified a school-based natural mentor during adolescence and individuals who identified no natural mentor. School-based natural mentorships are systematically more common among students from economically advantaged families and neighborhoods. Additionally, identifying no mentor at all is more common among students from homes and communities with less educational attainment. In Figure 2, we depict the likelihood of identifying a school-based natural mentor using binned averages of our SES measure and predicted values from a bivariate logistic regression of mentorship on SES. Students with SES values 1 SD above the median are 40% more likely to report a school-based natural mentor compared to students with SES scores about 1 SD below the median (17.5% and 12.5% predicted likelihood, respectively).

Finally, we find a slightly higher rate of exposure to school-based natural mentors for female students (15.4%) compared to males (14.8%), although there are more pronounced selection patterns by student gender across specific types of school personnel. Students who identify a teacher or counselor as a mentor are more likely to be female (61%), while students who identify a coach as their mentor are more likely to be male (71%). These patterns likely reflect the different gender compositions of teachers and counselors versus coaches and aligns with prior research on the salience of similar backgrounds (Ensher and Murphy, 1997; Stanton-Salazar and Urso Spina, 2003). Women comprise 41% of nominated K-12 teacher/counselor mentors, but only 13% of K-12 coach/athletic director mentors. Overall, male school personnel are more likely to be identified as natural mentors (59%) and are more likely to develop crossgender natural mentorships with female students (36% of all relationships with male mentors) compared to female personnel with male students (24% of all relationships with female mentors).

**II.c. What are the Characteristics of School-Based Natural Mentorships?**

Respondents characterized natural mentorships with school personnel as long-lasting, close relationships where mentors guide, advise, and encourage student development and self-realization, with an emphasis on academic attainment and life skills. Those who identified a school-based mentor reported that the relationship was important in their lives for more than five years on average. In fact, 80% reported that the relationship remained actively important in their lives in Wave III when respondents were between 18 and 26 years old. This highlights how K-12 mentors develop relationships that extend well beyond their formal roles within classrooms, counseling offices, and sports fields.

As we report in Table 2 Panel A, the hallmark behaviors of school-based mentors were providing guidance and sharing wisdom. Typical student descriptions of school-based mentors' actions included "gave me good advice," "taught me things," "gave me direction in life," "helped me stay out of trouble," and "helped me grow up." We also find that school-based mentors were frequently thought of as role models compared to other types of natural mentors.[8] Perhaps unsurprisingly, school-based natural mentors were unlikely to provide tangible help (e.g., financial support, transportation, fixing things) and were also rarely seen by their mentees as taking on the role of a family figure or friend.

We also find school-based mentors routinely participated in the developmental and academic lives of mentees. Table 2 Panel B shows that 64% of students with school-based mentors reported their mentors helped them with developmental outcomes. These outcomes

[8] Typical responses from this category included mention of the mentor being looked up to by the mentee, setting an example, being someone the mentee wants to be like, inspiring the mentee, and providing a positive influence.

tended to focus on self-development such as bringing out true qualities and strong moral character as well as life-development such as providing direction and helping students to make important decisions.[9] School-based natural mentors, particularly teachers and counselors, were also well positioned to engage with students' academic lives and inform their decisions regarding post-secondary education although only 34% of students identified this as a major domain of mentoring. K-12 mentors seldom inserted themselves into the religious, financial, and home lives of their mentees.[10] Although research suggests that informal mentorship is tailored to student-specific needs and can have broad-based benefits, these descriptive patterns motivate our focus on the relationship between school-based natural mentoring and students' academic achievement and attainment.

## III. Econometric Approach

Both prior research (Christensen et al., 2019; DuBois and Silverthorn, 2005a; Raposa et al., 2018; Zimmerman et al., 2005) and our descriptive analyses show that students from advantaged backgrounds are more likely to have a school-based natural mentor. Thus, a simple comparison of educational outcomes for students with and without school-based mentors would be biased upwards. In the absence of a randomized experiment, the empirical challenge is to address selection bias using non-experimental methods. This motivates our use of difference-in-differences TWFE models to remove all observed and unobserved time-invariant individual

[9] Typical responses from the developmental domain include the following: helped bring out true qualities; increased self-esteem; improved moral character; helped set goals; instilled confidence; gave direction; encouraged the right priorities; provided support during life crises; changed life; unspecified reference to decision-making or decisions.
[10] Supplemental analyses examining variation in the behavior and mentoring domains of K-12 mentors across student race reveal few consistent patterns or meaningful differences. We do find noticeable contrasts between the reported behaviors and mentoring domains of low-SES (bottom quintile) verses high-SES (top quintile) students. Low-SES students were more likely to report receiving practical, tangible help from their mentors and less likely to report receiving guidance, advice, and shared wisdom. Low-SES students were also more likely to report mentoring activities related to finances and money issues as well as school and college.

characteristics – e.g. family background, socioeconomic status, aptitude, race/ethnicity, gender, and native language – that influence the probability of engaging in these relationships.

**III.a. Two-Way Fixed Effects**

Our difference-in-differences TWFE model exploits variation within students over time to examine the relationship between school-based mentors and student outcomes in high school. We construct our analytic sample using student-level panel data with annual K-12 outcomes which we restrict to include only two groups of students: treated students who identified a school-based mentor starting after their freshman year and never-treated students who report never having had a natural mentoring relationship during adolescence. Specifically, we fit the following model:

$$y_{it} = \beta NM_{it}^{Sch} + \varphi_i + \gamma_t + \lambda_g + \varepsilon_{it} \tag{1}$$

where $y_{it}$ represents a high school transcript outcome for student *i* in year *t*. The first term on the right-hand side, $NM_{it}^{Sch}$, is an indicator variable which takes a value of 1 in all periods *t* in which a student had a K-12 natural mentor, and 0 otherwise. Individual student FEs ($\varphi_i$) focus our comparisons within students over time, while year ($\gamma_t$) and grade ($\lambda_g$) FEs control for any idiosyncratic shocks over time and across grade levels.[11]

Recent literature illustrates how TWFE models such as equation (1) can be biased in the presence of treatment effect heterogeneity (De Chaisemartin and D'Haultfoeuille, 2020; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2020). The specific nature of our data and empirical tests, however, suggest that this is unlikely to be a substantial concern within the

[11] In all models, grade FEs are based on a count of the number of years a student has been in high school.

context of our analysis given that the majority of our sample observations are from never-treated students (74%) and our short panel minimizes the potential for bias due to variance-weights (Goodman-Bacon, 2021). We estimate $\beta$ using the doubly robust approach developed by Callaway and Sant'Anna (2020) and Sant'Anna and Zhao (2020). This estimator removes any potential bias created by negative weights by estimating individual group-time average treatment effects (ATE), where the comparison group is restricted to never-treated units, and constructs a weighted average of these ATEs to produce a summary effect estimate for $\beta$. For all outcomes, we estimate cluster robust standard errors at the student level.

The identifying variation for our TWFE model comes from the differential timing of when students meet their natural mentors. The model assumes no other shocks to students' outcomes were concurrent with the timing of when they met their natural mentors, and that outcomes for students who report never having a mentor provide a valid counterfactual for the trends we would have otherwise observed among students with K-12 mentors. One concern might be that the timing of mentorship formation is endogenous, occurring at the same time as a change in students' effort or orientation towards school. However, prior evidence suggests the natural mentor formation process is a common phase of adolescent development for most students rather than a response to specific life events or other time-varying shocks (Rhodes, 2020). Among a sample of eleventh graders who identified having a very important non-parental adult in their lives, only 23% reported that these relationships arose when they were experiencing a significant event (Beam et al., 2002). Furthermore, for those students who were experiencing an important life event, nearly all of them reported navigating a negative experience like family or personal problems. This is consistent with findings from the broader social support literature that documents how adolescents are more likely to seek out help from a non-familial adult when they are experiencing major stressors (e.g. emotional problems, parental unemployment, parental separation, and family illness or death) (Hagler, Raposa, & Rhodes, 2019). These types of negative time-varying shocks at the onset of natural mentorship formation would, if anything, cause our estimates to understate any positive effects of natural mentorships on students' outcomes.

**III.b. Pair Fixed Effects**

Our analyses that employ individual student fixed effects are limited to short-run high school transcript outcomes for which we have panel data. We examine both short-run and long-run outcomes by fitting three separate models comparing outcomes between pairs of students who are twins (fraternal and identical), best-friends, or romantic partners.[12] We focus our discussion here on our preferred twins FE model, which removes many of the same fixed characteristics as our within-student TWFE model (family background, socioeconomic status, race/ethnicity, native language, etc.). We maximize our sample by including all pairs of twins, regardless of their natural mentor type. We then fit the following pair FE specification:

$$y_{ij} = \beta NM_{ij}^{Sch} + \sum \delta_z NM_{ij}^{z} + \gamma_j + \varepsilon_{ij} \tag{2}$$

where $y_{ij}$ represents a given long-run outcome of interest for student $i$ in pair $j$. We include FEs for each pair, $\gamma_j$, an indicator for having a school-based natural mentor, $NM_{ijb}^{Sch}$, and two indicators for students identifying the two other categories of non-K-12 natural mentors that we describe above, $NM_{ij}^{z}$. The coefficient, $\beta$, captures our parameter of interest: the relationship

[12] As shown in Appendix Figure B4, these three groups are largely distinct from one another with relatively few students included in more than one type of peer-pairing.

between a given outcome and having a school-based natural mentor relative to no mentor. For short-run outcomes that we observe in each year of high school, we estimate the model using panel data with repeated outcomes and include year and grade FEs similar to equation 1. This allows for more direct comparisons across our TWFE and pair FE models for high school outcomes.[13] We add birth cohort FEs in models with best-friend and romantic partner pairs to account for any age-related differences in long-run outcomes. We report cluster robust standard errors clustering on student pairs (i.e. twins, best friends, romantic partners) for both short- and long-run outcomes.[14]

A further advantage of Add Health's school-based sampling design is that all twin, best-friend, and romantic partner pairs in our data attended the same high school together. Thus, our pair FE models effectively include school FE. This is important given that school environments such as total enrollment, class size, resources, and school culture, might shape students' opportunities to develop natural mentoring relationships with school personnel.

The key identifying assumption of the twins FE models is that exposure to a school-based natural mentor is approximately random within twin pairs due, for example, to idiosyncrasies of the adults they meet in school. One might be concerned that our twins FE models (or other pair FE models) are biased due to selection based on individual student personality characteristics. We address these concerns by conducting balances tests, conditional on twins FE, using a total of nine individual and personality characteristics that vary within twins and are all significant predictors of school-based mentorship formation as shown in Table 1 and prior literature

[13] We drop freshman year outcomes to minimize attenuation bias resulting from most students meeting their natural mentor by sophomore year. About 1-in-3 students with school-based mentors had met their mentor in freshman year whereas nearly 75% had met their mentor by the end of sophomore year.

[14] We also estimate cluster robust standard errors for short-run outcomes by clustering at the individual student level to account for repeated observations. This alternative approach results in even more precise standard errors. Thus, we adopt the more conservative approach of clustering at the pair level for short-run outcomes.

(Erickson et al., 2009; Zimmerman et al., 2005; DuBois and Karcher, 2013).[15] These measures include: 1) whether a student is male; 2) whether the student has a learning difference; 3) the natural log of birthweight; 4) a five-point Likert-scale rating of physical health of an adolescent by his/her primary caretaker; 5) a count of the total number of friend nominations a student received as a proxy for extroversion and interpersonal skills; 6) a five-point Likert-scale rating of the physical attractiveness of an adolescent by an Add Health interviewer; 7) a five-point Likert-scale rating of the personality attractiveness of an adolescent by an Add Heath interviewer; 8) the number of sports teams and clubs a student was a member of in high school as a measure of engagement in extra-curricular activities; and 9) a five-point self-reported Likert-scale rating of the adolescents' closeness with their mother.

In Table 3, we report coefficients from a model in which we regress these student characteristics on an indicator for having a school-based natural mentor, conditional on twins FE. If exposure to school-based natural mentors is quasi-random within twin pairs, we would not expect to see any differences, on average, across these measures for twins with and without school-based mentors. Estimated coefficients for these nine predictors are uniformly small and statistically insignificant, and a joint F-test fails to reject the null of no joint predictive value ($p$=.87). We also demonstrate below that our primary results are robust to the inclusion of these additional covariates.

Our best-friend and romantic-partner FE models serve to further test the robustness of our student and twins FE estimators to potential selection on unobserved fixed and time-varying individual characteristics that influence students' likelihood of exposure to school-based natural

[15] The one exception is log birthweight ($p$=.16) which we retain given the large body of evidence documenting the consequential impacts of birthweight on human capital development (Black Devereux, & Salvanes al., 2007; Figlio et al., 2014).

mentors. A large literature documents the ways in which adolescents' social networks and relationships reflect selection on shared phenotypes, interests, personalities, and backgrounds (Shin and Ryan, 2014). We exploit this homophily and assortative matching that occurs among best-friend and romantic partner pairs. These models control for unobserved sets of shared fixed and time-varying individual characteristics – such as ability, preferences, and interests – that are shared within self-selected social pairs but may not be shared between twins. They are advantageous complements to our student and twins FE models because best-friend and romantic partner pairs remove selection based on unobserved time-varying personality characteristics during the same period when students are forming natural mentoring relationships in high school. Joint F-tests from conditional balance tests for the nine individual and personality characteristics described above also fail to reject the null for both best-friend and romantic-partner pairs (see Table 3).

**III.c. School Fixed Effects and Covariates**

Our pair FE models benefit from the ability to control for unobserved student characteristics, but have limited statistical power and do not take advantage of the nationally representative nature of the full dataset. As a complement and further robustness test, we fit models using the entire Add Health sample that employ high school FE and include a rich set of control variables. These models include direct controls for a wide range of observable time-varying and time-invariant student characteristics which prior research and descriptive analyses suggest influence students' likelihood of participating in a natural mentoring relationship. However, these models rely on stronger and less credible conditional independence assumptions than our student and pair FE models. The key advantage of this approach is that it allows us to test for potential heterogeneity across student characteristics that do not vary within twins and provides far greater precision than our pair FE models.

Our sample consists of all students, irrespective of their natural mentor type, who have valid data for the rich set of covariates we include. We fit the following model:

$$y_{isb} = \alpha + \beta NM_{isb}^{Sch} + \sum \delta_z NM_{isb}^{z} + \eta S_{isb} + \tau P_{isb} + \varphi C_{isb} + \pi_s + \theta_b + \varepsilon_{isb} \qquad (3)$$

where $y_{isb}$ represents a student-level outcome of interest for student $i$ in school $s$ in birth cohort $b$. This model replaces the pair FEs from equation 2 with school FEs ($\pi_s$), birth year FE ($\theta_b$), and adds covariate vectors for student ($S_{isb}$), parent ($P_{isb}$), and census-tract ($C_{isb}$) characteristics. Student variables include race, gender, interaction terms for race and gender, age, disability status, log birthweight, overall health, whether or not a student was born in the U.S., ever moved/relocated homes, ever separated from a caregiver, lived in a two-parent biological household, number of times nominated as a friend, interviewer perceptions of student physical and personality attractiveness, extracurricular participation, and maternal closeness. Parent variables include race, age, smoker status, disability status, educational attainment of the primary care provider and the primary care provider's partner, and whether or not the primary care provider was born in the U.S. Census-tract variables include population, population density, median household income, racial demographics, and the share of the population that is unemployed, without a high school diploma by the age of 25, without a college degree by the age of 25, owns their occupied homes, and receive welfare. For short-run academic outcomes which are measured annually, we include grade FEs and cluster our standard errors at the student level.

## IV. Findings

## IV.a. High School Academic Outcomes

Across all five of our fixed effect specifications, we find consistent evidence of a strong positive effect of having a K-12 natural mentor on students' short-run academic outcomes. This common pattern of results emerges despite each model using substantially different analytic samples, identifying variation, and identifying assumptions. Table 4 column 1 contains estimates from our preferred within-student TWFE specification, columns 2-4 contain estimates from our peer FE models (twins, best-friends, and romantic partners), and column 5 contains estimates from our school FE specification. Estimated effects on annual GPA range between 0.06 to 0.48 GPA points – a 2% to 20% increase compared to students who identify no mentor – and are significant across all models. Our most conservative estimate suggests that having a school-based mentor for all 4 years of high school raises GPA by 0.24 points or roughly the difference between a C+ and a B-.

We find striking evidence of the effect of school-based mentorships on the rate of annual course failure, with decreases ranging between 2.1 to 3.4 percentage points – a 22% to 35% reduction in the rate of course failure compared to unmentored students. Results are significant across all specifications with the exception of our least well powered romantic partners FE model. Complementing this reduction in failure rates, we also find that school-based mentors increase the number of course credits earned by 0.17 to 0.33 credits per year – a 3% to 5% increase. These estimates are similar in magnitude and significant in most models. Our most conservative estimate suggests that having a K-12 mentor for all 4 years of high school results in earning at least one additional semester-length credit.

## IV.b. Long-Run Academic Outcomes

We find compelling evidence of the effect of having a K-12 natural mentor on students' long-run human capital development. In Table 5 columns 1-4, we report the parameter estimates for having a K-12 natural mentor from our pair FE and school FE specifications. Compared to unmentored students, we estimate that having a natural mentor teacher, counselor, or coach increases the likelihood a mentee attends college by between 12 to 26 percentage points – a 19% to 46% increase. Estimates are significant across all but our least well-powered model (romantic partners FEs). Evidence on the role of K-12 mentors in helping students attend more selective colleges is mixed. All estimates are positive, while those from our twins FE models are near zero and those from our best friends, romantic partners, and school FE models are of sizable magnitude, 9 percentage points or greater.

Estimates for the effect of natural mentoring on total years of educational attainment are uniformly significant and strikingly similar in magnitude across models. These results reveal that K-12 mentoring increased educational attainment by between 0.62 and 0.93 years. These estimates imply that school-based mentorships raise the mean present value of lifetime earnings for high school freshmen by \$60,600 to \$92,400.[16] We next examine which specific level of students' higher education trajectories appears to drive this relationship with attainment. In Figure 3, we report estimates from a flexible, non-parametric modeling approach commonly called a distribution regression (Chernozhukov et al., 2013). Using our twins model, we estimate relationships at discrete levels of educational attainment and find that school-based natural mentors appear to benefit students most at the margin of supporting them to enroll in college. Our estimates from twin fixed effects models suggest a positive but very imprecisely estimated effect on 4-year college completion ($\beta$=0.068, s.e.=0.061).

[16] These values assume a 9% annual rate of return per additional year of education (Gunderson and Oreopoulos, 2020) and a median undiscounted lifetime earnings of \$1,037,000 in 2009 dollars (Carnevale et al., 2011).

The finding that K-12 natural mentors benefit students most on the margin of enrolling in college rather than completing a post-secondary degree is supported by theory and prior evidence. The college enrollment process occurs while students are still in high school and are more likely to benefit from regular interactions and support from their K-12 natural mentor (Rhodes, 2020). This pattern is also consistent with Gershensen et al. (2022) who find that same-race teachers benefit Black students by raising college matriculation rates in two-year programs, but not college completion. Although the Add Health data do not contain information to distinguish two-year and four-year degree programs, the uniformly positive but imprecise estimates on college selectivity suggest that K-12 natural mentors' effects are not entirely driven by supporting students to enroll in less-selective two-year degree programs.

**IV.c. Robustness Checks**

*Omitted Variable Bias*

We examine whether estimates from our twins, best-friend, and romantic-partner FE models are sensitive to the inclusion of the nine individual and personality characteristics we described above (See Appendix Tables B5 and B6). Across pair FE models and outcomes, we find that our estimates remain largely unchanged when these additional controls are added to the models. Estimates from twins FE models are very slightly attenuated, while estimates from best-friend FE models increase slightly, and estimates from romantic-partner models move only to a small degree in both directions. These findings suggest that unobserved time-varying student characteristics are unlikely to pose a substantial threat to our twins FE estimates.

We next conduct a bounding exercise developed by Oster (2019) to gauge the degree of selection bias on unobservables that would be needed to drive our estimates to zero. We implement the Oster procedure by first calculating the degree of selection explained by the large set of controls we add to our school FE model.[17] We then calculate $\delta$, the degree of selection on unobservables as a proportion of selection on observables, that would need to exist to negate our positive estimates (i.e. make $\beta = 0$) at a given $R^2$. We follow Oster's recommendation to set $R^2$ at $R_{max}$, which she estimates as $1.3*\tilde{R}$, where $\tilde{R}$ is the $R^2$ from our preferred school FE model with the full set of controls. We also calculate $\beta^*$, a lower bound estimate of our effects again assuming $R_{max}$= $1.3*\tilde{R}$ and setting selection on unobservables to be equal in magnitude to selection on observables ($\delta$=1).

The results from these bounding exercises displayed in Appendix Table B7 suggest that selection on unobservables would need to be two to three times as large as selection on observables to negate our findings. As shown in column 4, estimates of $\delta$ range from 1.6 to 3.1, with the largest estimate being the degree of selection necessary to negate our primary long-run finding of effects on college attendance. Lower bound estimates of $\beta^*$ are attenuated by between 30% and 50% but maintain their sign and remain economically meaningful in magnitude. For example, we estimate a 9.4 percentage point increase in the probability of attending college as a lower bound effect of a K-12 natural mentor. These sensitivity analyses suggest that positive selection into K-12 natural mentoring relationships based on unobserved characteristics is unlikely to account for the effects we find.

*Recall Bias*

Finally, we explore whether retrospective reporting of natural mentor relationships might bias our results. For example, our pair FE and school FE models would be biased upwards if

[17] Our school FE model is best suited for this sensitivity analysis for two reasons. It is our only model where we include a set of observable characteristics as controls, which is necessary to estimate selection on observables. It is also the only model where it is feasible to follow Oster's recommendations for estimating $R_{max}$ given that she suggests that the $R^2$ from a twin fixed effects itself proxies for $R_{max}$ – the upper bound estimate of potential variation that could be explained if all omitted variables were accounted for.

respondents who experience more academic or labor market success in life are more likely to view their relationships with school personnel as meaningful mentorships. Our within-student TWFE models are less sensitive to this type of recall bias as these models compare outcomes within students over time.

If recall bias exists, we might expect it to be stronger for older respondents who are farther removed from high school and who have more revealed information about their life outcomes. We test this by generating indicators for younger cohorts (Wave III ages 18-22.5) and older cohorts (Wave III ages 22.5-26) and interacting treatment with the younger cohort indicator as well as including it as a main effect.[18] This interaction tests for any differential effects across younger and older cohorts. Encouragingly, we find very few statistically significant differences across younger and older cohorts as shown in Appendix Tables B8 and B9. For those models where the interaction is significant, it uniformly suggests that effects are even slightly larger for the younger cohort, a pattern that is inconsistent with recall bias.

**IV.d. Heterogeneity**

*Student Characteristics*

The natural mentor literature offers theoretical rationales for why these relationships might dually compensate for a lack of access to resources for some students while also complementing resources among others. We use our most well-powered specification, the school FE model, to examine how different facets of student identity might moderate the effect of school-based mentorships on students' human capital development. Although we've shown that estimates from the school FE model are quite similar to our preferred TWFE and twins FE

[18] We re-estimate our within-student TWFE models separately for each sub-sample because current packages do not support interaction terms.

models, we view these results as exploratory rather than confirmatory given the stronger assumptions of the model and the large number of tests we conduct. We find some evidence consistent with the compensatory hypothesis based on students' SES, but our results largely suggest that students of different races and genders – and the intersections of these characteristics – benefit similarly from school-based mentorship.

We test for heterogeneity across SES by interacting our indicator for having a K-12 natural mentor in equation 3 with our composite measure of SES (results in Appendix Table B9). We find that having a K-12 mentor is a stronger predictor of reductions in course failure rates for students from lower SES backgrounds. As shown in Figure 4, the associated failure rate reduction among lower-SES students (1 SD below the median) is 3.9 percentage points, almost twice the 2.0 percentage point reduction associated with higher-SES students (1 SD above the median). For students' long-run educational attainment, our results suggest that school-based natural mentors are associated with a 16.7 percentage point increase in the probability of attending college for low-SES students and an 11.5 percentage point increase for high-SES students (Appendix Table B10). This translates to a 31% higher college-going rate among lower-SES students and a 14% higher rate among higher-SES students, suggesting a compensatory effect of mentorship. These findings are consistent with descriptive patterns of mentoring activities where low-SES students are more likely to report their mentor supported them in school and in pursuing college relative to students from average-SES and high-SES backgrounds. Estimates across other outcomes display similar patterns, but are smaller and statistically insignificant.

Estimates disaggregated by race and gender reveal few significant or systematic differences in our findings across the intersections of student characteristics we explore. The key exception is that school-based mentors appear particularly beneficial for Asian male students, the group for which we find the highest rate of reported K-12 mentoring relationships. As reported in Appendix Table B11, having a K-12 mentor is associated with a full letter grade improvement in GPA and 1.23 additional years of education for Asian male students.

*Mentor Type*

We next explore how school-based mentor effect estimates vary by mentor type. We can disaggregate our school-based natural mentor measure into two groups based on the Add Health data: teachers/counselors and coaches/athletic directors. We again leverage our highest-powered model using school FE given the small fraction of mentors who are coaches. As shown in Appendix Table B12, estimates for teachers/counselor and coach effects are quite similar across both short-run and long-run academic outcomes although less precisely estimated. For example, we estimate effects on GPA of 0.21 and 0.20 points for teachers/counselors and coaches, respectively. We find even slightly larger estimates for mentor coaches on students' long-run academic outcomes. Point estimates for effects on college going are 12.3 and 17.6 percentage points for teachers/counselors and coaches, respectively. These results add further evidence that the effects we find are not narrowly capturing the classroom-based effects that effective teachers have on their students.

**IV.e. External Validity**

Here we complement the range of internal validity tests presented above with a discussion of the external validity of our sample and research design. A key tradeoff of leveraging panel data that tracks adolescents into adulthood is that it requires us to rely on a sample of students that attended high school in the 1990s. Our pair fixed effects identification strategy also relies on a subset of students who are twins or report being in requited best-friend or romantic relationships. Although we demonstrate in Appendix Table B1 that these subgroups are broadly representative of the sample population as a whole, our estimates are only identified based on pairs that include one student with a K-12 natural mentor and another with a different type of mentor or no mentor at all.

We follow the recommendations of Miller, Shenhav, and Grosz (2021) to increase transparency about identifying variation in pair fixed effect designs by reporting the sample size of pairs that provide identifying variation and comparing their characteristics to pairs that do not contribute to our estimates. As show in Appendix Table B13, 202 students in our sample of 1,183 twins provide identifying variation. For best friends, the number is 340 out of 1,050 and for romantic partners, 112 out of 444. We next provide descriptive differences in the mean characteristics of FE pairs that do and do not provide identifying variation across each group as well as p-values from a difference in means test.

We find very few consistent differences across individual student characteristics with the exception of students being slightly more likely to be Latinx and less likely to be white across all three pair groups. Twins, and to a lesser degree best friends, that contribute to our estimates do appear to come from families that are more advantaged in terms of education levels, household income, and health. There are almost no meaningful differences based on census tract characteristics that emerge from these comparisons. These findings suggest that although our sets of student pairs are broadly similar to non-pairs, there does exist some degree of differences within pair groups among those that do and no not contribute to our identification. These non-random differences in identifying vs. non-identifying pairs along with the earlier decade in which our pair samples attended high school suggest that we should use caution when generalizing our findings to more recent cohorts and broader groups of students. **IV.f. School-level Correlates of School-Based Natural Mentor Relationships**

The importance of natural mentorships raises the question, "What can schools do to promote these relationships?" We explore this question by first documenting that the rate of K-12 natural mentorship varies considerably across schools. As shown in Figure 5, we estimate that schools at the 10th percentile of the distribution had 10% of students reporting a K-12 natural mentor while schools at the 90th percentile of the distribution had students reporting mentorships at more than double that rate. Next, we examine the degree to which school organizational and environmental factors predict students' likelihood of reporting a K-12 natural mentor versus no natural mentor at all using a linear probability model. In Table 6, we report results from simple bivariate regressions as well as a joint multivariate model. In both specifications we include our extensive set of student, parent, and neighborhood characteristics from our school FE models.

Our results reveal two significant school-level predictors of the likelihood students form school-based natural mentorships. First, schools in which students have a stronger sense of collective belonging have higher rates of school-based mentorships.[19] We estimate that a one SD increase in school-level peer average sense of belonging is associated with a 2 percentage point increase in the probability of having a K-12 natural mentor – an 13% increase. Second, smaller average class sizes predict higher rates of natural mentorship in schools. Our estimates suggest that for every ten fewer students in a classroom, on average, the probability a student forms a school-based natural mentorship increases by 2 percentage points. We do not find evidence that the formation of natural mentorships is correlated with other school environmental features such as the number of sports teams or clubs, urbanicity, or the average tenure of teachers in a school.

---

[19] We construct our belonging measure as a jackknife school-level mean of students' self-reported perceptions of belonging in school-based on responses to 7 survey items (see Appendix D for details).

## V. Conclusion

Schools serve as a cornerstone institution in society, generating substantial benefits for both individual students and the general public. Inside schools, students develop academic skills and content knowledge that have large returns in the labor market. Classroom learning, however, is not the only benefit schools provide. They also serve as social institutions where students interact with adults on a daily basis. Our paper highlights how these interactions can lead to the development of naturally occurring mentoring relationships with teachers, counselors, and coaches that extend well beyond when students leave their classrooms or team.

We find consistent evidence that having a school-based natural mentor increases academic performance and attainment across models based on disparate assumptions, identifying variation, and samples. Effects of school-based natural mentors on students' high-school grades are quite similar to those found in a recent field experiment evaluating the impacts of a year-long mentoring program for German youth from lower-socioeconomic backgrounds (Resnjanskij et al., 2021).[20] Our estimates on college enrollment are comparable or even larger than those found in studies of high-quality preschools (Gray-Lobe et al., 2021), no excuses charter schools (Angrist et al., 2016), double-dose remedial courses (Cortes et al., 2015; Ozek, 2021), and class size reductions (Chetty et al., 2011).

In contrast, we find substantially larger effects than prior research examining the long-run impacts of teachers (Chetty et al., 2014; Jackson, 2018; Petek and Pope, 2023).[21] Our research suggests that estimates of teachers' long-run impacts may understate their full effects in two

[20] As shown in Table 4, our estimated effect on annual GPA range from 0.056 to 0.478 points, with a standard deviation of 0.92. Resnjanskij et al. (2021) report an effect of 0.294 standard deviations.

[21] Petek and Pope (2021) find that a one standard deviation increase in elementary school teacher value-added to behavioral outcomes increases high school GPA by 0.013 points. Chetty et al. (2014) find that a one standard deviation increase in elementary and middle school teacher value-added to test scores increases college attendance by 0.86 percentage points. Jackson (2018) finds that a one standard deviation increase in 9th grade teacher value-added to behavioral outcomes increases GPA in 12th grade by 0.021 points.

ways. Prior research document teachers' long-run effects that operate exclusively through their value-added to academic or behavioral skills in a single year. Our study suggests that some teachers impact students' long-run outcomes through relationships that are sustained over time and through broader effects on their social capital, aspirations, and life decisions. This further reinforces growing evidence that teachers' impacts on students' long-run outcomes operate in large part through effects on students that are not captured by standardized tests (Jackson, 2018; Petek and Pope, 2023). Our findings most closely parallel the effects of same-race teachers on students' long-run outcomes such as college going which are similar in magnitude and likely operate to a large degree through role model effects (Gershenson et al., 2022).

We also document meaningful variation in the share of students reporting these relationships across schools, with mentorships occurring in some schools more than twice as often as others. This wide variation suggests that not all students have equal opportunity to develop school-based mentoring relationships. In particular, Black and Latinx students as well as low-SES students are meaningfully less likely to report having a school-based natural mentor. While this may be explained by the presence of other more impactful mentors in their lives, we expect that the lack of representation of these groups among school personnel contributes to the patterns we find. These findings add further motivation for efforts to recruit and retain a more diverse teacher workforce that can better represent and relate to the students they serve. They also highlight the importance of more effectively preparing all teachers with the cultural competencies to form meaningful relationships with students from diverse backgrounds.

Our findings point to several potential areas of future study. We need to better understand the characteristics of school-based natural mentors and whether some teachers, counselors, and coaches are more likely to serve in this capacity than others. We also lack a real understanding of the specific pathways through which school-based natural mentors support student mentees. Finally, we should explore how teacher training and school organizational practices might be leveraged to expand equitable access to these relationships. Our exploratory findings suggest that schools might promote these relationships by creating more opportunities for students to have multiple, sustained interactions with school personnel in small-group settings and by engendering school environments where all students feel a sense of belonging.

## References


Angrist, Joshua D., Sarah R. Cohodes, Susan M. Dynarski, Parag A. Pathak, and Christopher R. Walters. 2016. Stand and deliver: Effects of Boston's charter high schools on college preparation, entry, and choice. *Journal of Labor Economics* 34(2), 275-318.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. "School quality and the gender gap in educational achievement." *American Economic Review* 106, no. 5 (2016): 289-95.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. "Family disadvantage and the gender gap in behavioral and educational outcomes." *American Economic Journal: Applied Economics* 11, no. 3 (2019): 338-81.

Beam, Margaret R., Chuansheng Chen, and Ellen Greenberger. 2002. The nature of adolescents' relationships with their "very important" nonparental adults. *American Journal of Community Psychology,* 30(2), 305-325.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. "From the cradle to the labor market? The effect of birth weight on adult outcomes." *The Quarterly Journal of Economics* 122, no. 1 (2007): 409-439.

Blazar, David. (2021). Teachers of Color, Culturally Responsive Teaching, and Student Outcomes: Experimental Evidence from the Random Assignment of Teachers to Classes. EdWorkingPaper No. 21-501. *Annenberg Institute for School Reform at Brown University*.

Callaway, Brantly, and Pedro H.C. Sant'Anna. 2020. Difference-in-differences with multiple time periods. *Journal of Econometrics.*

Carnevale, Anthony. P., Ban Cheah, and Stephen J. Rose. 2011. The College Payoff. The Georgetown University Center on Education and the Workforce.

Chen, Ping, and Kathleen M. Harris. 2020. Guidelines for analyzing Add Health data. Carolina Population Center, University of North Carolina at Chapel Hill, 1-53.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2013. Inference on counterfactual distributions. *Econometrica* 81(6), 2205–2268.f

Chetty, Raj, John N. Friedman, Nahaniel Hilger, Emmanuel Saez, Diane W. Schanzenbach, and Danny Yagan. 2011. How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics* 126(4), 1593-1660.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review,* 104(9), 2633-79.

Christensen, Kirsten M., Elizabeth B. Raposa, Matthew A. Hagler, Lance Erickson, and Jean E. Rhodes. 2019. Role of athletic coach mentors in promoting youth academic success: Evidence from the add health national longitudinal study. *Applied Developmental Science,* 1-11.

Cortes, Kalena E., Joshua S. Goodman, and Takako Nomi. 2015. Intensive math instruction and educational attainment long-run impacts of double-dose algebra. *Journal of Human Resources,* 50(1), 108-158.

Currie, Janet, and Mark Stabile. "Child mental health and human capital accumulation: the case of ADHD." *Journal of health economics* 25, no. 6 (2006): 1094-1118.

Currie, Janet, and Erdal Tekin. "Understanding the cycle childhood maltreatment and future crime." *Journal of Human Resources* 47, no. 2 (2012): 509-549.

De Chaisemartin, Clément, and Xavier D'Haultfoeuille. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review* 110(9), 2964-96.

Dee, Thomas, S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of economics and statistics*, *86*(1), 195-210.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111-34.

Deutsch, Nancy L., Victoria A. Mauer, Haley E. Johnson, Anita A. Grabowska, and Miriam R. Arbeit. 2020. "[My counselor] knows stuff about me, but [my natural mentor] actually knows me": Distinguishing characteristics of youth's natural mentoring relationships. *Children and Youth Services Review* 111, 104879.

DuBois, David L., and Michael J. Karcher, eds. 2013. *Handbook of youth mentoring*. Thousand Oaks, CA: Sage Publications.

DuBois, David L., and Naida Silverthorn. 2005a. Natural mentoring relationships and adolescent health: Evidence from a national study. *American Journal of Public Health* 95(3), 518-524.

DuBois, David L., and Naida Silverthorn. 2005b. Characteristics of natural mentoring relationships and adolescent adjustment: Evidence from a national study. *Journal of Primary Prevention* 26(2), 69-92.

Duncan, Greg J., Johanne Boisjoly, and Kathleen Mullan Harris. 2001. Sibling, peer, neighbor, and schoolmate correlations as indicators of the importance of context for adolescent development. *Demography* 38(3), 437-447.

Elder, Glen H., Jr., Pavalko, Eliza K., and Clipp, Elizabeth C. 1993. *Working with archival data: Studying lives.* Newbury Park, CA: Sage.

Ensher, Ellen A., and Susan E. Murphy. 1997. Effects of race, gender, perceived similarity, and contact on mentor relationships. *Journal of Vocational Behavior* 50(3), 460-481.

Erickson, Lance D., Steve McDonald, and Glen H. Elder Jr. 2009. Informal mentors and education: Complementary or compensatory resources? *Sociology of Education* 82(4), 344-367.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. "The effects of poor neonatal health on children's cognitive development." *American Economic Review* 104, no. 12 (2014): 3921-55.

Fruiht, Veronica M., and Laura Wray-Lake. 2013. The role of mentor type and timing in predicting educational attainment. *Journal of Youth and Adolescence* 42(9), 1459-1472.

Gershenson, Seth, Cassandra MD Hart, Joshua Hyman, Constance A. Lindsay, and Nicholas W. Papageorge. (2022). The long-run impacts of same-race teachers. *American Economic Journal: Economic Policy* 14(4): 300-342.

Goodman-Bacon, Andrew. 2021. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Gowdy, Grace, Daniel P. Miller, and Renée Spencer. 2020. Expanding and deepening our understanding of which young people are most likely to have an informal mentor. *Children and Youth Services Review*, 108, 104490.

Gunderson, Morley, and Philip Oreopolous. 2020. Returns to education in developed countries. In *The Economics of Education*, ed Steve Bradley, and Colin Green. Academic Press.

Granovetter, Mark S. 1973. The strength of weak ties. *American journal of sociology* 78(6), 1360-1380.

Gray-Lobe, Guthrie, Parag A. Pathak, and Christopher R. Walters. 2021. The long-term effects of universal preschool in Boston NBER Working Paper w28756, National Bureau of Economic Research, Cambridge, MA.

Grossman, Jean Baldwin, and Joseph P. Tierney. 1998. Does mentoring work? An impact study of the Big Brothers Big Sisters program. *Evaluation review* 22(3), 403-426.

Hagler, M., Raposa, E.B. and Rhodes, J., 2019. Psychosocial profiles of youth who acquire a natural mentor during a school year. *Applied Developmental Science*, *23*(2), pp.144-152.

Harris, Kathleen Mullen, Carolyn Tucker Halpern, Brett C. Haberstick, and Andrew Smolen. 2013. The national longitudinal study of adolescent health (Add Health) sibling pairs data. *Twin Research and Human Genetics* 16(1), 391-398.

Herrera, Carla, Jean Baldwin Grossman, Tina J. Kauh, and Jennifer McMaken. 2011. Mentoring in schools: An impact study of Big Brothers Big Sisters school-based mentoring. *Child development* 82(1), 346-361.

Hurd, Noelle M., Jamie Albright, Audrey Wittrup, Andrea Negrete, and Janelle Billingsley. 2018. Appraisal support from natural mentors, self-worth, and psychological distress: Examining the experiences of underrepresented students transitioning through college. *Journal of Youth and Adolescence* 47*,* 1100-1112.

Hurd, Noelle M., Bernadette Sánchez, Marc A. Zimmerman, and Cleopatra H. Caldwell. 2012. Natural mentors, racial identity, and educational attainment among African American adolescents: Exploring pathways to success. *Child Development* 83(4), 1196-1212.

Hurd, Noelle M., and Robert M. Sellers. 2013. Black adolescents' relationships with natural mentors: Associations with academic engagement via social and emotional development. *Cultural Diversity and Ethnic Minority Psychology* 19(1), 76-85.

Hurd, Noelle M., Joseph S. Tan, and Emily L. Loeb. 2016. Natural mentoring relationships and the adjustment to college among underrepresented students. *American Journal of Community Psychology* 57*,* 330-341.

Hurd, Noelle M., and Marc. A. Zimmerman. 2014. An analysis of natural mentoring relationship profiles and associations with mentees' mental health: Considering links via support from important others. *American Journal of Community Psychology* 53(1-2), 25-36.

Hussar, Bill, Jijun Zhang, Sarah Hein, Ke Wang, Ashley Roberts, Jiashan Cui, Mary Smith, Farrah Bullock Mann, Amy Barner, and Rita Dilig. 2020. The condition of education. NCES 2020-144, National Center for Education Statistics, Washington DC.

Jacinto, Alberto, and Seth Gershenson. 2021. The intergenerational transmission of teaching. *American Educational Research Journal* 58(3), 635-672.

Jackson, Kirabo C. 2018. What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy*. 126(5), 2072-2107.

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk. 2020. The formation of prosociality: causal evidence on the role of social environment. *Journal of Political Economy* 128(2), 434-467.

Miranda-Chan, Thomas, Veronica Fruiht, Valeska Dubon, and Laura Wray-Lake. 2016. The functions and longitudinal outcomes of adolescents' naturally occurring mentorships. *American Journal of Community Psychology* 57(1-2), 47-59.

Mulhern, Christine. 2020. Beyond teachers: Estimating individual guidance counselors' effects on educational attainment. Institute of Education Sciences, Harvard University, Cambridge, MA.

Oreopoulos, Philip, and Kjell G. Salvanes. 2011. Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives* 25(1), 159-84.

Ozek, Umut. 2021. The effects of middle school remediation on postsecondary success: Regression discontinuity evidence from Florida. *Journal of Public Economics*, 203, 104518

Petek, Nathan, and Nolan Pope. 2023. The multidimensional impact of teachers on students. *Journal of Political Economy*, 131(4), 1057.

Raposa, Elizabeth B., Lance D. Erickson, Matthew Hagler, and Jean E. Rhodes. 2018. How economic disadvantage affects the availability and nature of mentoring relationships during the transition to adulthood. *American Journal of Community Psychology* 61(1-2), 191-203.

Resnjanskij, Sven., Jens Ruhose, Simon Wiederhold, and Ludger Woessmann. 2021. Can mentoring alleviate family disadvantage in adolescence? A field experiment to improve labor-market prospects. CESifo Working Paper 8870, Munich, Germany.

Rhodes, Jean E. 2005. A model of youth mentoring. *Handbook of Youth Mentoring* 30-43.

Rhodes, Jean E., and David L. DuBois. 2006. Understanding and facilitating the youth mentoring movement. *Social policy report*, 20(3), 1-20.

Rhodes, J.E., 2020. *Older and Wiser*. Harvard University Press.

Rhodes, Jean E., Lori Ebert, and Karla Fischer. 1992. Natural mentors: An overlooked resource in the social networks of young, African American mothers. *American Journal of Community Psychology* 20, 445-461.

Sánchez, Bernadette, Patricia Esparza, and Yarí Colón. 2008. Natural mentoring under the microscope: An investigation of mentoring relationships and Latino adolescents' academic performance. *Journal of Community Psychology* 36(4), 468-482.

Shin, Huiyoung, and Allison M. Ryan. 2014. Early adolescent friendships and academic adjustment: Examining selection and influence processes with longitudinal social network analysis. *Developmental Psychology* 50(11), 2462.

Stanton-Salazar, Ricardo D., and Sanford M. Dornbusch. 1995. Social capital and the reproduction of inequality: Information networks among Mexican-origin high school students. *Sociology of education* 116-135.

Stanton-Salazar, Ricardo D. and Stephanie Urso Spina. 2003. Informal mentors and role models in the lives of urban mexican-origin adolescents. *Anthropology and Education Quarterly* 34, 231–54.

Van Dam, L., D. Smit, B. Wildschut, S.J.T. Branje, J.E. Rhodes, M. Assink, and G.J.J. Stams. 2018. Does natural mentoring matter? A multilevel meta-analysis on the association between natural mentoring and youth outcomes. *American Journal of Community Psychology* 62(1-2), 203-220.

Zimmerman, Marc A., Jeffrey B. Bingenheimer, and D.E. Behrendt. 2005. Natural mentoring relationships. In *Handbook of Youth Mentoring,* eds David L. DuBois and Michael J. Karcher. Thousand Oaks, CA: Sage Publications.

# Tables

Table 1. Sample Characteristics

| | Add Health | HS NM | No NM | p-value |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel A: Student Characteristics | | | | |
| Asian/Pac. Isl. | 0.03 | 0.04 | 0.04 | 0.61 |
| Black | 0.15 | 0.11 | 0.17 | 0.00 |
| Latinx | 0.12 | 0.10 | 0.15 | 0.00 |
| white | 0.65 | 0.70 | 0.59 | 0.00 |
| Male | 0.51 | 0.48 | 0.53 | 0.00 |
| Age in 1994 (years) | 15.96 | 15.51 | 15.99 | 0.00 |
| Students with disabilities | 0.15 | 0.09 | 0.18 | 0.00 |
| English spoken at home | 0.92 | 0.94 | 0.89 | 0.01 |
| US Born | 0.92 | 0.91 | 0.90 | 0.44 |
| General health (0-4) | 3.11 | 3.23 | 3.02 | 0.00 |
| Log(Birthweight) | 4.76 | 4.77 | 4.76 | 0.16 |
| Times nominated as a best friend | 0.52 | 0.59 | 0.52 | 0.08 |
| Number of clubs and sports | 1.55 | 2.01 | 1.34 | 0.00 |
| Considered attractive, physically | 0.49 | 0.50 | 0.46 | 0.05 |
| Considered attractive, personality | 0.49 | 0.53 | 0.44 | 0.00 |
| Close relationship with mother | 0.84 | 0.86 | 0.82 | 0.02 |
| Always lived in same home | 0.22 | 0.23 | 0.22 | 0.52 |
| Biological father present | 0.87 | 0.91 | 0.87 | 0.00 |
| Panel B: Parent Characteristics | | | | |
| Age in 1994 (years) | 41.42 | 41.47 | 40.98 | 0.07 |
| Disabled | 0.06 | 0.05 | 0.07 | 0.02 |
| US Born | 0.88 | 0.89 | 0.84 | 0.01 |
| Recently accepted welfare | 0.10 | 0.07 | 0.12 | 0.00 |
| Neither parent has HS diploma | 0.15 | 0.09 | 0.20 | 0.00 |
| HS diploma highest deg. earned | 0.24 | 0.22 | 0.28 | 0.00 |
| Attended some college | 0.31 | 0.31 | 0.29 | 0.29 |
| Highest degree is bachelors | 0.17 | 0.18 | 0.15 | 0.02 |
| Graduate schooling | 0.13 | 0.20 | 0.09 | 0.00 |
| Household income in 1994 ($) | 45,190 | 49,595 | 40,427 | 0.00 |
| At least very good health | 0.48 | 0.56 | 0.43 | 0.00 |
| Smoker in household | 0.48 | 0.44 | 0.50 | 0.00 |
| Panel C: Census Tract Characteristics | | | | |
| Population | 5,633 | 5,892 | 5,578 | 0.13 |
| Asian/Pac. Isl. | 0.03 | 0.03 | 0.03 | 0.77 |
| Black | 0.14 | 0.12 | 0.15 | 0.00 |
| Latinx | 0.08 | 0.07 | 0.10 | 0.03 |
| white | 0.79 | 0.82 | 0.78 | 0.00 |
| Pop. without HS diploma by 25 | 0.27 | 0.25 | 0.30 | 0.00 |

| | | | | |
|---|---|---|---|---|
| Pop. without coll. degree by 25 | 0.23 | 0.24 | 0.21 | 0.00 |
| Household income ($) | 29,704 | 31,852 | 27,803 | 0.00 |
| Pop. on welfare | 0.09 | 0.08 | 0.10 | 0.00 |
| Owner occupied dwelling | 0.68 | 0.69 | 0.66 | 0.01 |
| Unemployment rate | 0.08 | 0.07 | 0.08 | 0.00 |
| n(students) | 18,924 | 2,185 | 3,702 | |

Note. Values represent portion of data unless otherwise noted. P-value compares students with a school-based mentor to those with no mentor. Add Health provided weights are used to achieve national representativeness.

Table 2. Percent of respondents who mention a behavior or domain when asked "what did your mentor do to help you?"

| | All school-based mentors | Teachers & counselors | Coaches & athletic directors | Rank among all mentor types (1-5) |
|---|---|---|---|---|
| Panel A: Behaviors of Mentor | | | | |
| Guidance, advice, shared wisdom | 69.5 | 70.0 | 66.7 | 1 |
| Emotional nurturance | 38.7 | 39.5 | 34.9 | 4 |
| Practical, tangible help (labor performed) | 3.9 | 3.9 | 4.0 | 5 |
| Like a parent, mother figure, father figure | 1.9 | 1.6 | 3.2 | 5 |
| Like a friend | 5.2 | 4.5 | 8.4 | 4 |
| Role model | 14.6 | 13.8 | 18.4 | 2 |
| Spend time together | 1.5 | 1.2 | 3.2 | 4 |
| Other | 2.0 | 2.0 | 1.7 | 4 |
| Panel B: Domains of Mentoring | | | | |
| Developmental outcomes (life & self) | 64.4 | 61.4 | 78.7 | 2 |
| Family and household | 2.5 | 2.5 | 2.3 | 5 |
| Religion | 2.3 | 1.8 | 5.1 | 5 |
| Finances, money issues | 1.5 | 1.7 | 0.7 | 5 |
| Work, job | 9.2 | 10.4 | 3.5 | 4 |
| School, college | 33.8 | 38.4 | 11.9 | 1 |
| Time together, leisure, sports, social | 11.4 | 8.3 | 25.8 | 2 |
| Other | 26.4 | 26.9 | 24.2 | 5 |
| n(students) | 2,185 | 1,761 | 424 | |

Note. Other categories include responses that identify a behavior and/or domain of mentoring which does not fit in any of the other categories. The ranking column provides the 1 to 5 ranking of how frequently a behavior or domain is mentioned about school-based mentors relative to other mentor types including family members, friends, non-familial adults, and mentors met after HS (1 being most frequent). See Appendix A for a description of the Add Health response coding scheme.

Table 3. Balance on select observables between twin, best-friend, and romantic partner pairs conditional on pair fixed effects

| | Variable Type | Twins | Best-Friends | Romantic Partners |
|---|---|---|---|---|
| Male | Binary | -0.041 | 0.371 | 0.007 |
| | | (0.036) | (0.278) | (0.031) |
| SWD status | Binary | -0.015 | -0.026 | -0.025 |
| | | (0.047) | (0.055) | (0.072) |
| Times nominated as a best friend | Count | 0.018 | -0.001 | -0.002 |
| | | (0.018) | (0.010) | (0.011) |
| Attractiveness - Physical | Likert | 0.006 | 0.013 | 0.030 |
| | | (0.020) | (0.021) | (0.027) |
| Attractiveness - Personality | Likert | 0.012 | -0.022 | 0.008 |
| | | (0.023) | (0.019) | (0.027) |
| Participation in sch. clubs & sports | Count | -0.004 | 0.004 | 0.011 |
| | | (0.006) | (0.007) | (0.011) |
| Closeness to mother | Likert | 0.007 | -0.036** | -0.018 |
| | | (0.020) | (0.018) | (0.028) |
| General health | Likert | -0.005 | -0.015 | 0.009 |
| | | (0.027) | (0.018) | (0.024) |
| Log(Birthweight) | Continuous | -0.041 | -0.104 | 0.051 |
| | | (0.116) | (0.099) | (0.144) |
| p-score of joint F-test | | 0.835 | 0.420 | 0.827 |
| n(students) | | 1,261 | 1,390 | 556 |

Note. * p<0.10, ** p<0.05, *** p<0.01. Column 1 describes the distribution/potential values for each variable. Columns 2-4 report the coefficients from a model where each characteristic is regressed on an indicator for having a school-based natural mentor conditional on pair FE. Columns 2, 3, and 4 condition on twins FE, best-friend FE, and romantic partner FE, respectively. SWD = Students with disabilities. Standard errors are clustered within twins, best friend, and romantic partner pairs.

Table 4. The relationship between school-based natural mentorship and students' short-run education outcomes

| | Mean of unmentored students | Within-student TWFE | Twin FE | Best friend FE | Romantic partner FE | School FE |
|---|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) | (5) |
| Panel A: Annual Cumulative GPA (0-4.0) | | | | | | |
| HS Mentor | 2.47 | 0.061*** | 0.265** | 0.200** | 0.478*** | 0.206*** |
| | | (0.021) | (0.107) | (0.083) | (0.151) | (0.028) |
| n(student-years) | | 14,325 | 2,872 | 3,406 | 1,384 | 16,904 |
| Panel B: Annual Percent of Courses Failed | | | | | | |
| HS Mentor | 9.56 | -2.081*** | -3.360* | -2.243* | -3.360 | -2.551*** |
| | | (0.589) | (1.773) | (1.251) | (2.953) | (0.496) |
| n(student-years) | | 14,325 | 2,872 | 3,406 | 1,384 | 16,904 |
| Panel C: Annual Year-long Courses Passed | | | | | | |
| HS Mentor | 5.82 | 0.166*** | 0.334** | 0.134 | 0.206 | 0.209*** |
| | | (0.056) | (0.151) | (0.103) | (0.259) | (0.042) |
| n(student-years) | | 14,441 | 2,901 | 3,425 | 1,388 | 17,013 |
| Periods included | | All | Post-Freshman Year | | | |
| Calendar year FE | | Yes | Yes | Yes | Yes | Yes |
| Grade FE | | Yes | Yes | Yes | Yes | Yes |
| Birth year FE | | | | Yes | Yes | Yes |
| Controls | | | | | | Yes |

Note. * p<0.10, ** p<0.05, *** p<0.01. TWFE = Two-way fixed effects. FE = fixed effects. Standard errors are clustered within individuals for the TWFE and school FE models and clustered within pairs for the twins, best friend, and romantic partner FE models. Controls include student, family, and neighborhood characteristics. Student variables include race, gender, an interaction term for race and gender, age at wave 1, SWD status, log of birthweight, general health assessment, whether a student was born in the US, the number of times a student was nominated as a close friend by others, extracurricular participation, and interviewer ratings of student physical and personality-based attractiveness. Family variables include primary caregiver's race, highest education attained, general health, age, US born status, household income in 1994, and whether or not a smoker lives in the household, English is the primary language at home, the biological father was ever present in childhood, and the present household includes both biological parents. Neighborhood variables are based on census tracts and include population, population density, the portion of the population that is white, Black, Asian/Pacific Islander, Latinx, earned a high school diploma by age 25, earned a college degree by age 25, receives welfare, owns the house they occupy, is unemployed, and average household income.

Table 5. The relationship between school-based natural mentorship and students' long-run academic outcomes

| | Mean of unmentored students | Twin FE | Best friend FE | Romantic partner FE | School FE |
|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) |
| Panel A: Attended College | | | | | |
| HS Mentor | 0.54 | 0.154** | 0.257*** | 0.118 | 0.134*** |
| | | (0.063) | (0.055) | (0.083) | (0.017) |
| n(students) | | 1,025 | 1,081 | 426 | 6,411 |
| Panel B: Attended a Selective College | | | | | |
| HS Mentor | 0.18 | 0.025 | 0.093* | 0.167 | 0.078*** |
| | | (0.057) | (0.054) | (0.108) | (0.017) |
| n(students) | | 1,025 | 1,081 | 426 | 6,411 |
| Panel C: Lifetime Educational Attainment in Years | | | | | |
| HS Mentor | 14.53 | 0.725** | 0.930*** | 0.842** | 0.622*** |
| | | (0.318) | (0.258) | (0.400) | (0.076) |
| n(students) | | 1,025 | 1,081 | 426 | 6,411 |
| n(FE pairs) | | 691 | 695 | 278 | |
| Birth year FE | | | Yes | Yes | Yes |
| Controls | | | | | Yes |

Note. * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered within pairs for the twins, best friend, and romantic partner FE models. We estimate heteroskedasticity robust standard errors for the school FE model. Controls include student, family, and neighborhood characteristics and the full list of variables can be found in Table 4. Long-run educational outcomes are measured in Wave IV (respondents age 24-32).

Table 6. School-level correlates of having a school-based natural mentor

| | Bivariate | Multivariate |
|---|---|---|
| Enrollment (100 students) | -0.001 | 0.000 |
| | (0.001) | (0.001) |
| Class size | -0.004*** | -0.002* |
| | (0.001) | (0.001) |
| Number of student clubs | -0.002 | 0.003 |
| | (0.003) | (0.003) |
| Number of student sports | -0.003 | 0.003 |
| | (0.011) | (0.013) |
| Student sense of belonging (SDs) | 0.021*** | 0.020*** |
| | (0.006) | (0.006) |
| Suburban setting | -0.010 | 0.000 |
| | (0.014) | (0.012) |
| Rural setting | -0.004 | 0.009 |
| | (0.014) | (0.016) |
| Observations | 6,799 | |

Note. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Suburban and rural setting indicators are relative to an urban default category. Belonging refers to the standardized jackknife mean of school level average responses to survey items asking about student perceptions of belonging/safety/trust at school, see Appendix D for more details. All models include the full set of student, family, and neighborhood characteristics described in Table 4.

Figures

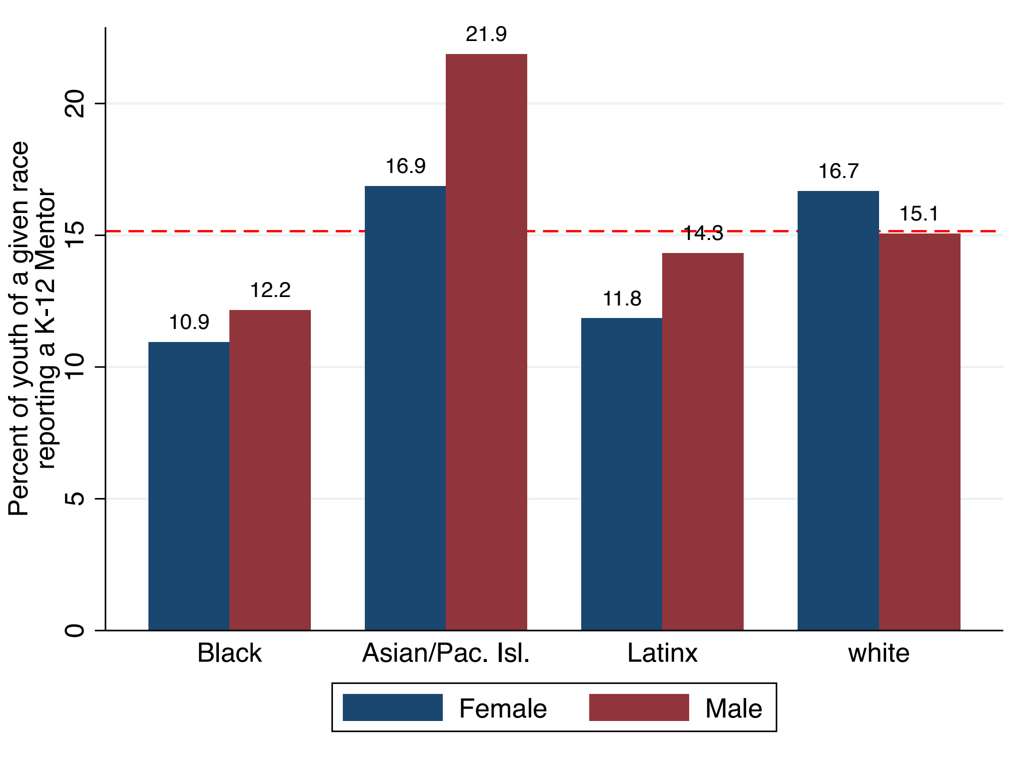



Figure 1. The frequency of school-based natural mentor relationships by race and gender

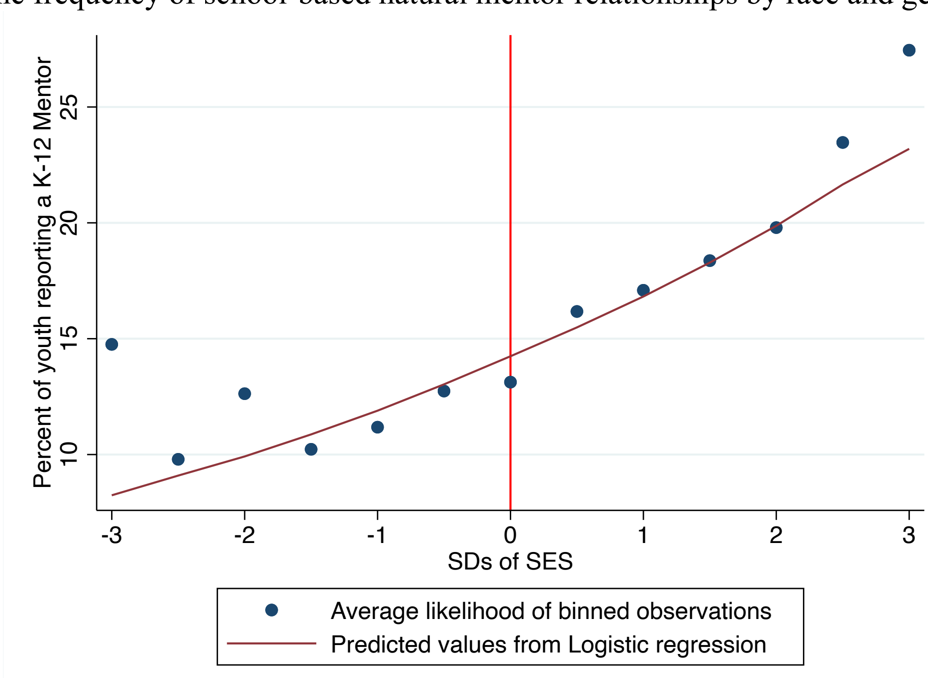



Figure 2. The likelihood of identifying a school-based natural mentor based on socioeconomic status

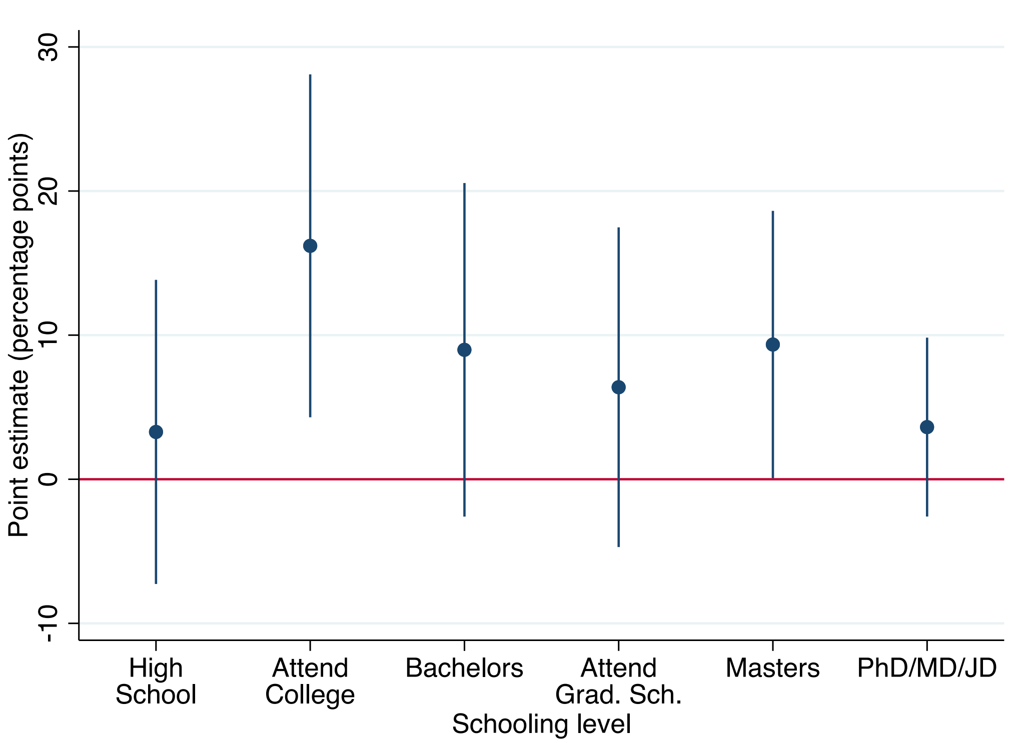
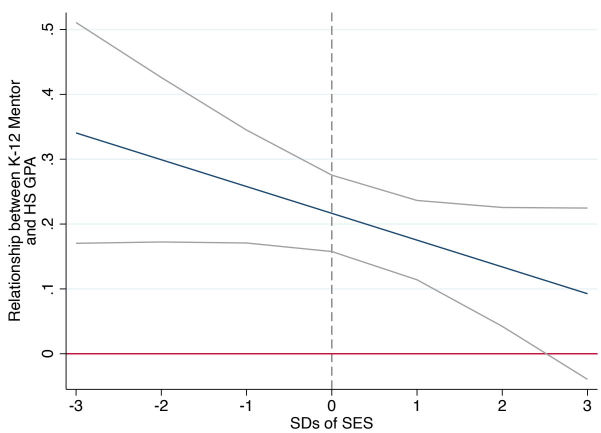
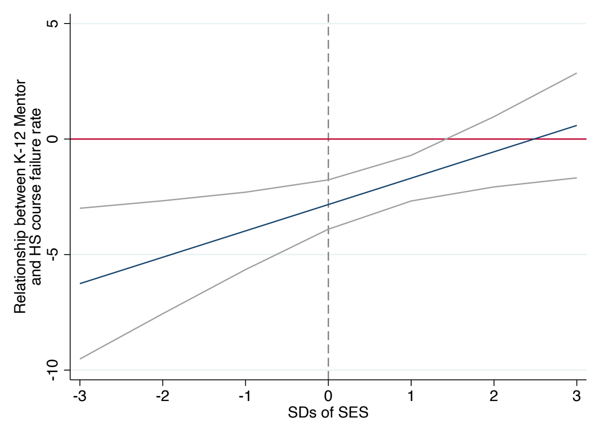



Figure 3. Point estimates with 95% CIs from a distribution regression of the association between identifying a school-based natural mentor and educational attainment

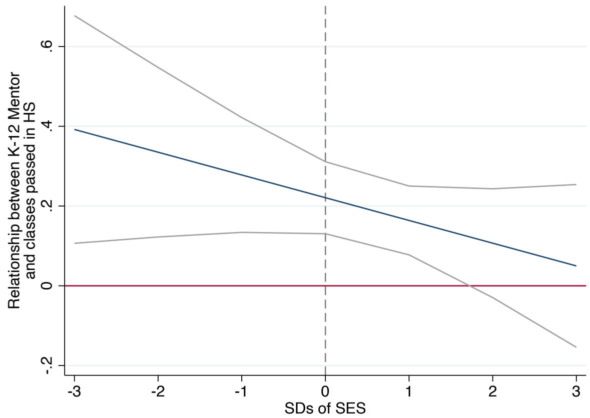
Panel A: Annual high school GPA



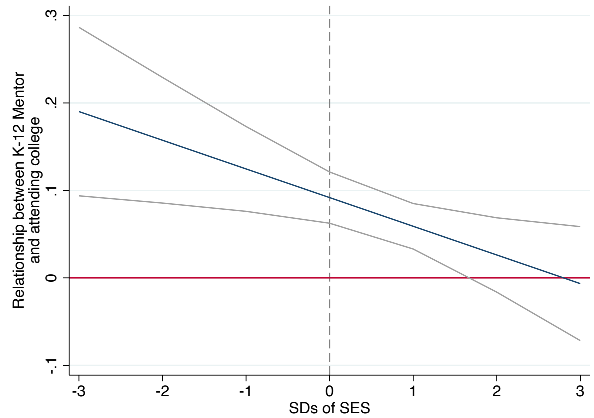

Panel B: Annual rate of course failure



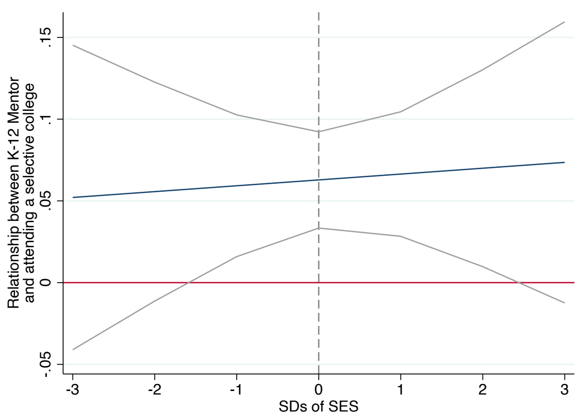
Panel C: Annual year-length courses passed




Panel D: College Attendance




Panel E: Attending a college with a selective admissions process




Panel F: Lifetime educational attainment

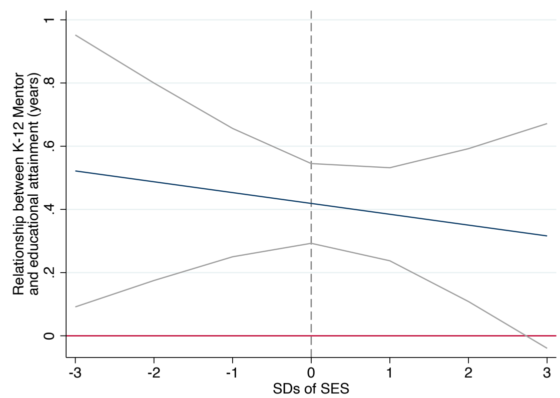



Figure 4. Linear estimates of heterogeneity in the relationship between having a school-based natural mentor and outcomes from school FE models.

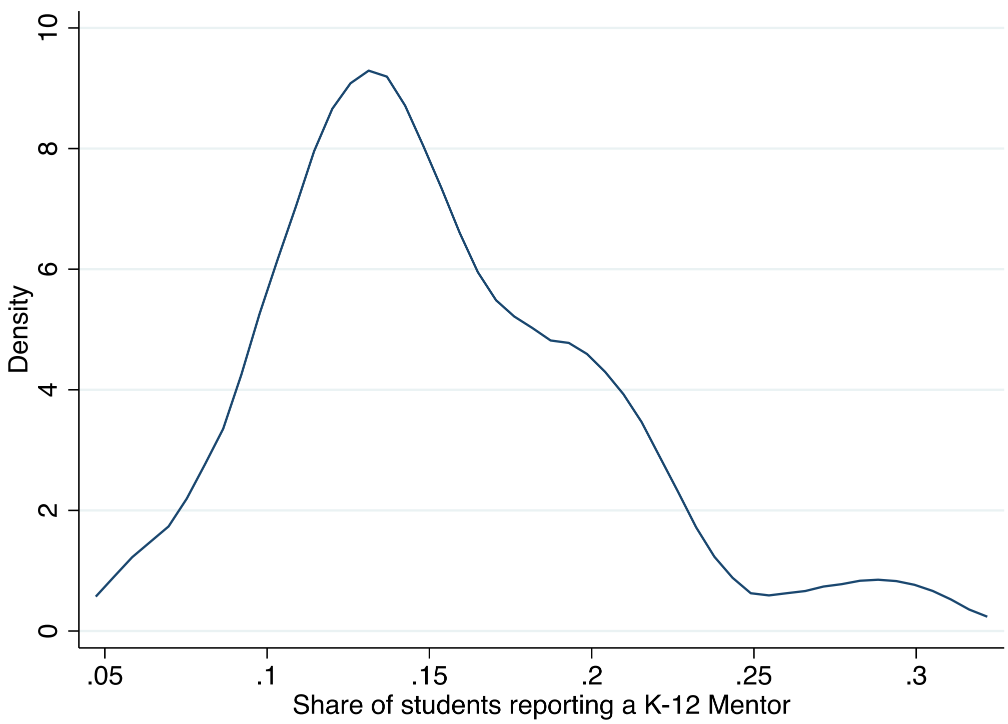



Figure 5. Kernel density of the portion of students in a school who identify a school-based mentor among schools with at least 10 respondents (n=85)

## Appendix A

The following excerpt is taken from the Add Health Wave III codebook:

At Wave III, Add Health respondents were asked if they had a mentor. Respondents who reported a mentor were asked about their mentor's functional role by responding to an open-ended question, "What did [your mentor] do to help you?" A coding scheme was inductively developed from their responses based on a group of randomly selected cases. All code development work was done using an approximate 10% random sample (N=1048) of responses.

Responses to the mentor item were divided into 12 sets of 1,000 and one set of 489 responded. Coding was done by two researchers and each one was given 550 responses, in which 100 of the same responses were given to both researchers to calculate kappas. This resulted in 10 percent of the cases being coded by both researchers, where were used to calculate inter-rater reliability. This method to calculate Kappa is reviewed in Elder, Pavalko, Clip (1993). Intermediate kappa's were calculated within each set of 100 using the 100 cases given to both researchers and for each category (variable). When all responses had been assigned codes, a pooled Kappa was calculated for all the items both researchers coded. Kappas ranged from .79 to .96 for [all behaviors and domains], indicating a high degree of coder agreement (Elder et al., 1993).

The responses were coded into categories that described the behavior of the mentor and the domain of mentoring. The categories were not mutually exclusive—an individual's response would be coded in more than one category. Each element of the response was coded for both behavior and domain.

Table A1. Kappa statistics for mentor behavior and domain codes

| Variable | Pooled Kappa | Variable | Pooled Kappa |
|---|---|---|---|
| MENTORA1 | 0.79 | MENTORB1 | 0.82 |
| MENTORA2 | 0.87 | MENTORB2 | 0.85 |
| MENTORA3 | 0.80 | MENTORB3 | 0.95 |
| MENTORA4 | 0.95 | MENTORB4 | 0.91 |
| MENTORA5 | 0.96 | MENTORB5 | 0.87 |
| MENTORA6 | 0.86 | MENTORB6 | 0.93 |
| MENTORA7 | 0.59 | MENTORB7 | 0.76 |
| MENTORA8 | 0.32 | MENTORB8 | 0.73 |

# Appendix B

Appendix Table B1. Characteristics for different peer-pair analytic samples

| | Twins | Non-twins | p-value | Friends | Non-Friends | p-value | Partners | Non-partners | p-value |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| **Panel A: Student Characteristics** | | | | | | | | | |
| Asian/Pac. Isl. | 0.03 | 0.03 | 0.00 | 0.09 | 0.03 | 0.00 | 0.08 | 0.03 | 0.17 |
| Black | 0.25 | 0.15 | 0.06 | 0.16 | 0.15 | 0.00 | 0.13 | 0.15 | 0.00 |
| Latinx | 0.16 | 0.12 | 0.27 | 0.13 | 0.13 | 0.00 | 0.16 | 0.12 | 0.78 |
| white | 0.53 | 0.65 | 0.01 | 0.58 | 0.64 | 0.00 | 0.59 | 0.65 | 0.00 |
| Male | 0.50 | 0.51 | 0.59 | 0.38 | 0.52 | 0.00 | 0.50 | 0.51 | 0.88 |
| Age in 1994 (years) | 16.03 | 15.97 | 0.38 | 16.12 | 15.98 | 0.39 | 16.56 | 15.95 | 0.00 |
| Students with disabilities | 0.14 | 0.15 | 0.56 | 0.07 | 0.15 | 0.00 | 0.06 | 0.15 | 0.00 |
| English spoken at home | 0.92 | 0.92 | 0.02 | 0.91 | 0.92 | 0.01 | 0.92 | 0.92 | 0.01 |
| US Born | 0.93 | 0.92 | 0.00 | 0.90 | 0.92 | 0.13 | 0.92 | 0.92 | 0.03 |
| General health (0-4) | 3.20 | 3.10 | 0.00 | 3.21 | 3.10 | 0.00 | 3.22 | 3.10 | 0.00 |
| Log(Birthweight) | 4.51 | 4.77 | 0.00 | 4.75 | 4.76 | 0.08 | 4.76 | 4.76 | 0.01 |
| Times nominated as a best friend | 0.46 | 0.51 | 0.00 | 2.41 | 0.39 | 0.00 | 2.35 | 0.48 | 0.00 |
| Number of clubs and sports | 1.50 | 1.54 | 0.10 | 1.99 | 1.52 | 0.00 | 1.96 | 1.54 | 0.03 |
| Considered attractive, physically | 0.51 | 0.49 | 0.02 | 0.52 | 0.49 | 0.00 | 0.60 | 0.49 | 0.00 |
| Considered attractive, personality | 0.50 | 0.49 | 0.14 | 0.53 | 0.49 | 0.01 | 0.57 | 0.49 | 0.00 |
| Close relationship with mother | 0.84 | 0.84 | 0.21 | 0.85 | 0.84 | 0.02 | 0.81 | 0.84 | 0.33 |
| Always lived in same home | 0.24 | 0.22 | 0.00 | 0.26 | 0.22 | 0.00 | 0.25 | 0.22 | 0.17 |
| Biological father present | 0.87 | 0.87 | 0.00 | 0.90 | 0.87 | 0.00 | 0.90 | 0.87 | 0.00 |
| **Panel B: Parent Characteristics** | | | | | | | | | |
| Age in 1994 (years) | 42.40 | 41.41 | 0.00 | 42.21 | 41.43 | 0.14 | 41.78 | 41.42 | 0.55 |
| Disabled | 0.06 | 0.06 | 0.03 | 0.04 | 0.06 | 0.00 | 0.06 | 0.06 | 0.50 |
| US Born | 0.86 | 0.88 | 0.05 | 0.83 | 0.88 | 0.67 | 0.85 | 0.88 | 0.24 |
| Recently accepted welfare | 0.10 | 0.10 | 0.08 | 0.06 | 0.10 | 0.00 | 0.06 | 0.10 | 0.00 |
| Neither parent has HS diploma | 0.16 | 0.15 | 0.59 | 0.12 | 0.15 | 0.00 | 0.11 | 0.15 | 0.00 |

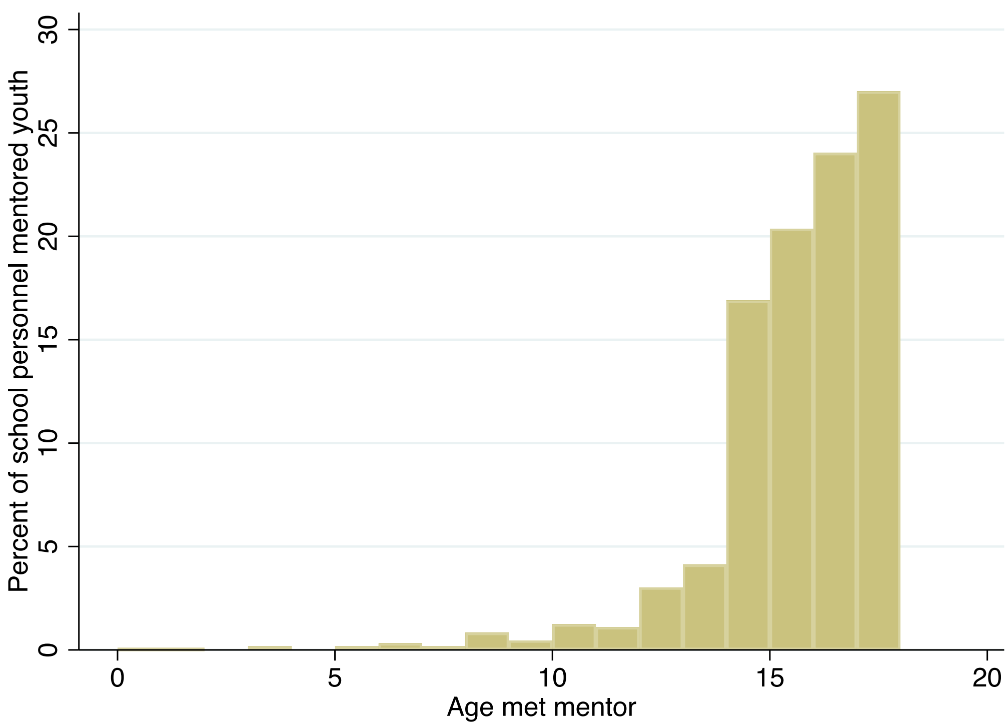
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| HS diploma highest deg. earned | 0.17 | 0.24 | 0.00 | 0.22 | 0.24 | 0.82 | 0.22 | 0.24 | 0.94 |
| Attended some college | 0.31 | 0.31 | 0.64 | 0.30 | 0.31 | 0.54 | 0.30 | 0.31 | 0.80 |
| Highest degree is bachelors | 0.21 | 0.17 | 0.00 | 0.19 | 0.17 | 0.18 | 0.23 | 0.17 | 0.00 |
| Graduate schooling | 0.14 | 0.13 | 0.02 | 0.17 | 0.13 | 0.00 | 0.15 | 0.13 | 0.81 |
| Household income in 1994 ($) | 47,446 | 45,187 | 0.06 | 49,251 | 44,817 | 0.02 | 50,620 | 45,032 | 0.05 |
| At least very good health | 0.49 | 0.48 | 0.09 | 0.51 | 0.48 | 0.00 | 0.48 | 0.48 | 0.21 |
| Smoker in household | 0.44 | 0.48 | 0.00 | 0.44 | 0.48 | 0.33 | 0.44 | 0.48 | 0.53 |
| Panel C: Census Tract Characteristics | | | | | | | | | |
| Population | 5823.85 | 5640.21 | 0.00 | 5284.77 | 5665.69 | 0.00 | 5213.24 | 5650.02 | 0.02 |
| Asian/Pac. Isl. | 0.04 | 0.03 | 0.00 | 0.06 | 0.03 | 0.00 | 0.06 | 0.03 | 0.06 |
| Black | 0.19 | 0.14 | 0.47 | 0.14 | 0.14 | 0.00 | 0.12 | 0.14 | 0.00 |
| Latinx | 0.11 | 0.08 | 0.22 | 0.10 | 0.08 | 0.01 | 0.11 | 0.08 | 0.82 |
| white | 0.72 | 0.79 | 0.63 | 0.75 | 0.79 | 0.00 | 0.77 | 0.79 | 0.00 |
| Pop. without HS diploma by 25 | 0.27 | 0.27 | 0.00 | 0.28 | 0.27 | 0.09 | 0.28 | 0.27 | 0.81 |
| Pop. without coll. degree by 25 | 0.24 | 0.23 | 0.00 | 0.22 | 0.23 | 0.27 | 0.22 | 0.23 | 0.13 |
| Household income ($) | 31,158 | 29,696 | 0.00 | 32,477 | 29,608 | 0.00 | 32,321 | 29,694 | 0.00 |
| Pop. on welfare | 0.09 | 0.09 | 0.00 | 0.09 | 0.09 | 0.00 | 0.08 | 0.09 | 0.00 |
| Pop. that owns dwelling | 0.67 | 0.68 | 0.01 | 0.72 | 0.67 | 0.00 | 0.72 | 0.68 | 0.00 |
| Unemployment rate | 0.08 | 0.08 | 0.02 | 0.07 | 0.08 | 0.01 | 0.07 | 0.08 | 0.00 |
| n(students) | 1,209 | 17,636 | | 1,384 | 17,551 | | 554 | 18,377 | |

Note. Values are shares unless otherwise noted. Add health weights are used to achieve national representativeness.

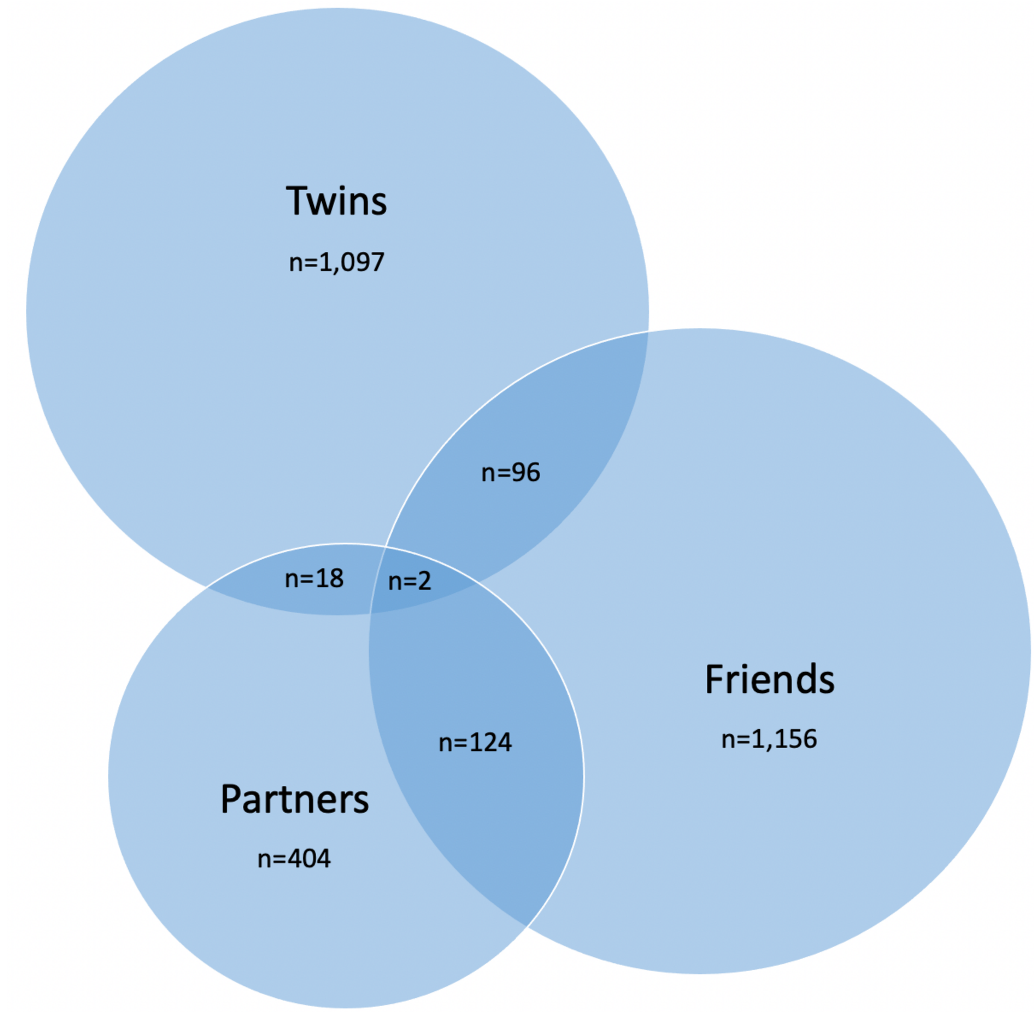
Appendix Table B2. Summary statistics for outcomes of interest

| Outcome | Outcome type | Mean | Count |
|---|---|---|---|
| Annual GPA (0-4.0) | Measured | 2.63 | 44,554 |
| Annual percent of courses failed | Measured | 7.53 | 44554 |
| Annual semester length classes passed | Measured | 5.96 | 44,871 |
| Years of Education | Self-reported | 14.99 | 14,796 |
| HS completion* | Self-reported | 0.83 | 14,800 |
| Some college* | Self-reported | 0.64 | 14,800 |
| College degree* | Self-reported | 0.30 | 14,800 |
| Grad school* | Self-reported | 0.08 | 14,800 |

Note. *Indicates binary variable. Measured outcomes are taken from official high school transcripts; Self-reported outcomes are based on Wave IV survey responses or Wave III when missing Wave IV. We achieve nationally representative estimates using Add Health provided weights.




Appendix Figure B3. Histogram of the age when a respondent reports meeting a school-based natural mentor



Appendix Figure B4. Visual representation of twins, best-friends, and romantic partners sample overlaps

Table B5. The relationship between school-based natural mentorship and students' short-run education outcomes conditional on controls

| | Mean of unmentored students | Twin FE | | Best friend FE | | Romantic partner FE | |
|---|---|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) | (5) | (6) |
| | | Panel A: Annual Cumulative GPA (0-4.0) | | | | | |
| HS Mentor | 2.47 | 0.265** | 0.223** | 0.200** | 0.215*** | 0.478*** | 0.425*** |
| | | (0.107) | (0.096) | (0.083) | (0.081) | (0.151) | (0.161) |
| n(student-years) | | 2872 | | 3406 | | 1384 | |
| | | Panel B: Annual Percent of Courses Failed | | | | | |
| HS Mentor | 9.56 | -3.360* | -2.683* | -2.243* | -2.443** | -3.360 | -2.613 |
| | | (1.773) | (1.528) | (1.251) | (1.243) | (2.953) | (3.085) |
| n(student-years) | | 2872 | | 3406 | | 1384 | |
| | | Panel C: Annual Year-long Courses Passed | | | | | |
| HS Mentor | 5.82 | 0.334** | 0.295** | 0.134 | 0.137 | 0.206 | 0.157 |
| | | (0.151) | (0.138) | (0.103) | (0.106) | (0.259) | (0.262) |
| n(student-years) | | 2901 | | 3425 | | 1388 | |
| Periods included | | | | Post-Freshman Year | | | |
| Calendar year FE | | Yes | Yes | Yes | Yes | Yes | Yes |
| Grade FE | | Yes | Yes | Yes | Yes | Yes | Yes |
| Birth year FE | | | | Yes | Yes | Yes | Yes |
| Controls | | | Yes | | Yes | | Yes |

Note. * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered within pairs for all specifications. Controls include student gender, SWD status, log of birthweight, times nominated as a best-friend, self-reported closeness to mother and general health, and a student's physical and personality attractiveness as determined by an Add Health interviewer.

Table B6. The relationship between school-based natural mentorship and students' long-run academic outcomes conditional on controls

| | Mean of unmentored students | Twin FE | | Best friend FE | | Romantic partner FE | |
|---|---|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) | (5) | (6) |
| | | | | Panel A: Attended College | | | |
| HS Mentor | 0.54 | 0.154** | 0.138** | 0.257*** | 0.256*** | 0.118 | 0.103 |
| | | (0.063) | (0.059) | (0.055) | (0.054) | (0.083) | (0.082) |
| n(students) | | 1,025 | | 1,081 | | 426 | |
| | | | | Panel B: Attended a Selective College | | | |
| HS Mentor | 0.18 | 0.025 | 0.016 | 0.093* | 0.102* | 0.167 | 0.146 |
| | | (0.057) | (0.056) | (0.054) | (0.053) | (0.108) | (0.102) |
| n(students) | | 1,025 | | 1,081 | | 426 | |
| | | | | Panel C: Lifetime Educational Attainment in Years | | | |
| HS Mentor | 14.53 | 0.725** | 0.613** | 0.930*** | 0.978*** | 0.842** | 0.802** |
| | | (0.318) | (0.312) | (0.258) | (0.258) | (0.400) | (0.391) |
| n(students) | | 1,025 | | 1,081 | | 426 | |
| Birth year FE | | | | Yes | Yes | Yes | Yes |
| Controls | | | Yes | | Yes | | Yes |

Note. * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered within pairs for all specifications. Controls include student gender, SWD status, log of birthweight, times nominated as a best-friend, self-reported closeness to mother and general health, and a student's physical and personality attractiveness as determined by an Add Health interviewer.

Table B7. Testing for sensitivity to selection on unobservables using Oster bounds

| Outcome | School FE | $R^2$ | $R$max | δ | β* |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Panel A: Short Run Outcomes | | | | | |
| Annual Cumulative GPA (0-4.0) | 0.206 | 0.301 | 0.391 | 2.962 | 0.139 |
| Annual Percent of Courses Failed | -2.551 | 0.165 | 0.215 | 2.512 | -1.573 |
| Annual Year-long Courses Passed | 0.166 | 0.173 | 0.225 | 1.616 | 0.082 |
| Panel B: Long Run Outcomes | | | | | |
| Attended College | 0.134 | 0.259 | 0.337 | 3.111 | 0.094 |
| Attended a Selective College | 0.078 | 0.238 | 0.310 | 2.017 | 0.041 |
| Lifetime Educational Attainment in Years | 0.622 | 0.372 | 0.484 | 2.623 | 0.396 |

Notes. Column 1 reports the coefficient from the School FE specifications in Tables 4 & 5. Column 2 reports R-squared from these specifications. Column 3 reports the $R$max value used to calculate Oster bounds which in all cases is 1.3*R-squared as suggested in Oster (2019). Column 4 reports delta, the magnitude of selection on unobservables relative to selection on observables required for a true treatment effect of 0. Column 5 reports β*, the treatment effect estimate under the assumption that the magnitude of selection on unobservables is the same as selection on observables.

Appendix Table B8. The relationship between natural mentorship and short-run academic outcomes for students from older and younger cohorts

| | Within-student TWFE | | Twin FE | Best friend FE | Romantic partner FE | School FE |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | | | Panel A: Annual GPA (0-4.0) | | | |
| HS Mentor | 0.067** | 0.039 | 0.351** | 0.038 | 0.440* | 0.220*** |
| | (0.030) | (0.041) | (0.168) | (0.108) | (0.234) | (0.041) |
| Young (age<median) | | | 0.309 | 0.135 | 0.162 | -0.038 |
| | | | (0.263) | (0.165) | (0.205) | (0.055) |
| Young*HS NM | | | -0.158 | 0.344** | 0.092 | -0.005 |
| | | | (0.220) | (0.166) | (0.311) | (0.055) |
| n(student-years) | 6,911 | 7,397 | 2,872 | 3,406 | 1,384 | 17,595 |
| | | | Panel B: Annual Percent of Courses Failed | | | |
| HS Mentor | -2.259*** | -1.333 | -3.003 | -0.281 | -1.722 | -2.298*** |
| | (0.775) | (1.349) | (2.486) | (1.602) | (5.093) | (0.746) |
| Young (age<median) | | | -4.398 | 0.657 | 0.546 | 2.039* |
| | | | (4.766) | (3.506) | (3.442) | (1.069) |
| Young*HS NM | | | -0.814 | -4.484* | -4.241 | -0.920 |
| | | | (3.603) | (2.567) | (5.906) | (1.006) |
| n(student-years) | 6,911 | 7,397 | 2,872 | 3,406 | 1,384 | 17,595 |
| | | | Panel C: Annual Year-length Courses Passed | | | |
| HS Mentor | 0.198*** | 0.148 | 0.444** | 0.008 | -0.024 | 0.155** |
| | (0.070) | (0.097) | (0.193) | (0.123) | (0.396) | (0.060) |
| Young (age<median) | | | 0.482 | -0.008 | -0.012 | -0.181** |
| | | | (0.368) | (0.256) | (0.377) | (0.089) |
| Young*HS NM | | | -0.200 | 0.280 | 0.547 | 0.134 |
| | | | (0.308) | (0.216) | (0.492) | (0.084) |
| n(student-years) | 6,973 | 7,447 | 2,901 | 3,425 | 1,388 | 17,713 |
| Sample | Younger | Older | All | All | All | All |
| Periods included | All | All | | Post-Freshman Year | | |
| Calendar year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Grade FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Birth year FE | | | | Yes | Yes | Yes |
| Controls | | | | | | Yes |

Note. TWFE = Two-way fixed effects. FE = fixed effects. Standard errors are clustered within individuals for the TWFE and school FE models and clustered within pairs for the twins, best friend, and romantic partner FE models. Controls include student, family, and neighborhood characteristics and the full list of variables can be found in Table 3. Columns 1 and 2 report coefficients for DiD models estimated on the sample students in the youngest 50 percentile and the oldest 50 percentile, respectively. The younger cohort consists of students below the median age.

Appendix Table B9. The relationship between natural mentorship and long-run academic outcomes for students from older and younger cohorts

| | Twin FE | Best friend FE | Romantic partner FE | School FE |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel A: Attended College | | | | |
| HS Mentor | -0.066 | 0.192** | 0.090 | 0.124*** |
| | (0.090) | (0.077) | (0.097) | (0.024) |
| Young (age<median) | -0.128** | -0.028 | 0.074 | 0.000 |
| | (0.064) | (0.110) | (0.145) | (0.032) |
| Young*HS NM | 0.375*** | 0.128 | 0.031 | 0.024 |
| | (0.123) | (0.112) | (0.169) | (0.033) |
| n(students) | 1,025 | 1,081 | 426 | 6,663 |
| Panel B: Attended a Selective College | | | | |
| HS Mentor | -0.076 | 0.015 | 0.210 | 0.078*** |
| | (0.091) | (0.075) | (0.134) | (0.026) |
| Young (age<median) | -0.188 | 0.041 | 0.239* | -0.032 |
| | (0.154) | (0.091) | (0.142) | (0.028) |
| Young*HS NM | 0.170 | 0.134 | -0.114 | 0.003 |
| | (0.117) | (0.108) | (0.209) | (0.033) |
| n(students) | 1,025 | 1,081 | 426 | 6,369 |
| Panel C: Lifetime Educational Attainment in Years | | | | |
| HS Mentor | 0.004 | 0.572 | 0.901** | 0.567*** |
| | (0.541) | (0.362) | (0.443) | (0.114) |
| Young (age<median) | 0.339 | 0.105 | 0.151 | 0.015 |
| | (0.627) | (0.494) | (0.623) | (0.142) |
| Young*HS NM | 1.224* | 0.658 | -0.166 | 0.133 |
| | (0.664) | (0.514) | (0.824) | (0.148) |
| n(students) | 1,025 | 1,081 | 426 | 6,366 |
| Birth Year FE | | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes |

Note. * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered within pairs for the twins, best friend, and romantic partner FE models. We estimate heteroskedasticity robust standard errors for the school FE model. Controls include student, family, and neighborhood characteristics and the full list of variables can be found in Table 4. The younger cohort consists of students below the median age.

Appendix Table B10. Heterogeneity tests for short- and long-run academic outcomes by SES

| | Annual HS GPA (0-4.0) | | | Annual HS Fail % | | | Annual year-long classes passed | | |
|---|---|---|---|---|---|---|---|---|---|
| **Panel A: High School Transcript Outcomes** | | | | | | | | | |
| K-12 Mentor | 0.215*** | 0.223*** | 0.250*** | -2.734*** | -2.962*** | -3.987*** | 0.224*** | 0.233*** | 0.297*** |
| | (0.028) | (0.030) | (0.066) | (0.497) | (0.546) | (1.292) | (0.042) | (0.046) | (0.110) |
| SES measure | | 0.092* | | | -2.503** | | | 0.226*** | |
| | | (0.051) | | | (1.082) | | | (0.084) | |
| K-12 Mentor*SES | | -0.036 | | | 0.968** | | | -0.044 | |
| | | (0.024) | | | (0.439) | | | (0.039) | |
| SES quintile 2 | | | 0.084 | | | -0.651 | | | 0.037 |
| | | | (0.087) | | | (1.598) | | | (0.136) |
| SES quintile 3 | | | -0.075 | | | 1.957 | | | -0.099 |
| | | | (0.083) | | | (1.501) | | | (0.129) |
| SES quintile 4 | | | -0.066 | | | 1.852 | | | -0.160 |
| | | | (0.078) | | | (1.387) | | | (0.120) |
| SES quintile 5 | | | -0.072 | | | 2.127 | | | -0.082 |
| | | | (0.074) | | | (1.324) | | | (0.117) |
| n(students) | | | 17,595 | | | | | 17,713 | |
| **Panel B: Long-run Educational Outcomes** | | | | | | | | | |
| | Attended College | | | Attended a Selective College | | | Educational Attainment in Years | | |
| K-12 Mentor | 0.136*** | 0.141*** | 0.194*** | 0.079*** | 0.076*** | 0.052 | 0.638*** | 0.629*** | 0.702*** |
| | (0.017) | (0.018) | (0.045) | (0.017) | (0.017) | (0.035) | (0.074) | (0.076) | (0.176) |
| SES measure | | 0.010*** | | | 0.073*** | | | 0.611*** | |
| | | (0.031) | | | (0.023) | | | (0.123) | |
| K-12 Mentor*SES | | -0.026** | | | 0.007 | | | -0.004 | |
| | | (0.013) | | | (0.014) | | | (0.063) | |
| SES quintile 2 | | | -0.049 | | | 0.039 | | | -0.083 |
| | | | (0.056) | | | (0.048) | | | (0.224) |
| SES quintile 3 | | | -0.037 | | | 0.023 | | | -0.038 |
| | | | (0.055) | | | (0.047) | | | (0.219) |

| | | | |
|---|---|---|---|
| SES quintile 4 | -0.096* | 0.011 | -0.116 |
| | (0.051) | (0.046) | (0.214) |
| SES quintile 5 | -0.077* | 0.052 | -0.066 |
| | (0.047) | (0.045) | (0.201) |
| n(students) | 6,663 | | 6,660 |
| Primary FE | | School FE | |
| Controls | | Yes | |

Note. *p<.1, **p<.05, ***p<0.1. We estimate heteroskedasticity robust standard errors. All models control for student, family, and neighborhood characteristics. The full list of control variables can be found in Table 4. All models are based on our school FE specification.

Appendix Table B11. Heterogeneity tests for educational outcomes at the intersection of race and gender

| | Panel A: High School Transcript Outcomes (annual measures) | | | | | |
|---|---|---|---|---|---|---|
| | GPA (0-4) | | Course Failure % | | Year-long classes passed | |
| K12 Mentor | 0.215*** | 0.171*** | -2.734*** | -1.387** | 0.224*** | 0.113** |
| | (0.028) | (0.041) | (0.497) | (0.648) | (0.042) | (0.057) |
| K12*white male | | 0.048 | | -1.767** | | 0.104 |
| | | (0.055) | | (0.810) | | (0.074) |
| K12*Black female | | 0.054 | | -1.049 | | 0.399*** |
| | | (0.100) | | (1.525) | | (0.148) |
| K12*Black male | | -0.004 | | -2.033 | | -0.029 |
| | | (0.104) | | (2.210) | | (0.187) |
| K12*Asian female | | -0.057 | | -0.303 | | 0.083 |
| | | (0.136) | | (1.735) | | (0.177) |
| K12*Asian male | | 0.351*** | | -6.097*** | | 0.408** |
| | | (0.131) | | (1.799) | | (0.168) |
| K12*Latina | | 0.087 | | -1.707 | | 0.134 |
| | | (0.089) | | (1.524) | | (0.143) |
| K12*Latino | | -0.030 | | -1.436 | | 0.189 |
| | | (0.099) | | (1.818) | | (0.145) |
| n(students) | | 17,595 | | | 17,713 | |

| | Panel B: Long-run Educational Outcomes | | | | | |
|---|---|---|---|---|---|---|
| | Attended College | | Attended a Selective College | | Educational Attainment (Yrs.) | |
| K12 Mentor | 0.136*** | 0.098*** | 0.079*** | 0.053** | 0.638*** | 0.512*** |
| | (0.017) | (0.024) | (0.017) | (0.027) | (0.074) | (0.113) |
| K12*white male | | 0.091*** | | 0.069* | | 0.229 |
| | | (0.032) | | (0.038) | | (0.149) |
| K12*Black female | | 0.038 | | 0.048 | | 0.014 |
| | | (0.046) | | (0.059) | | (0.236) |
| K12*Black male | | 0.022 | | -0.076 | | -0.128 |
| | | (0.062) | | (0.057) | | (0.262) |

| | | | |
|---|---|---|---|
| K12*Asian female | -0.184** | -0.114 | -0.155 |
| | (0.085) | (0.097) | (0.442) |
| K12*Asian male | 0.121* | 0.112 | 0.610 |
| | (0.067) | (0.109) | (0.431) |
| K12*Latina | -0.006 | -0.053 | 0.081 |
| | (0.057) | (0.061) | (0.262) |
| K12*Latino | 0.065 | 0.065 | 0.252 |
| | (0.062) | (0.057) | (0.244) |
| n(students) | 6,663 | | 6,660 |
| Model | School FE | | |
| Controls | Yes | | |

Note. *p<.1, **p<.05, ***p<0.01. We estimate heteroskedasticity robust standard errors. All models control for student, family, and neighborhood characteristics. The full list of control variables can be found in Table 4.

Table B12. The relationship between school-based natural mentor teachers and coaches and students' short- and long-run academic outcomes

| | (1) | (2) | (3) |
|---|---|---|---|
| | Panel A: Short-run Outcomes | | |
| | Annual GPA (0-4.0) | Annual Percent of Courses Failed | Annual Year-length Courses Passed |
| Teacher | 0.209*** | -2.308*** | 0.235*** |
| | (0.031) | (0.531) | (0.045) |
| Coach | 0.196*** | -3.555*** | 0.118* |
| | (0.043) | (0.654) | (0.062) |
| n(students) | 16,904 | 16,904 | 17,013 |
| | Panel B: Long-run Outcomes | | |
| | Attended college | Attended a selective college | Lifetime ed. attainment in years |
| Teacher | 0.123*** | 0.0605*** | 0.579*** |
| | (0.0181) | (0.0184) | (0.0810) |
| Coach | 0.176*** | 0.146*** | 0.805*** |
| | (0.0253) | (0.0325) | (0.122) |
| n(students) | 6,414 | 6,414 | 6,411 |

Note. * p<0.10, ** p<0.05, *** p<0.01. We estimate heteroskedasticity robust standard errors. All Panel A models include control variables and separate fixed effects for the school, calendar year, student grade, and birth year. All Panel B models include controls and separate fixed effects for students' school and birth year. Controls include student gender, SWD status, log of birthweight, times nominated as a best-friend, self-reported closeness to mother and general health, and a student's physical and personality attractiveness as determined by an Add Health interviewer.

Table B13. Sample characteristics for FE pairs that do and do not provide identifying variation for estimating school-based natural mentor effects

| | Twin FE | | | Best friend FE | | | Romantic partner FE | | |
|---|---|---|---|---|---|---|---|---|---|
| | FE Pairs that Provide Identification | FE Pairs that do NOT Provide Identification | P-value of Difference | FE Pairs that Provide Identification | FE Pairs that do NOT Provide Identification | P-value of Difference | FE Pairs that Provide Identification | FE Pairs that do NOT Provide Identification | P-value of Difference |
| Panel A: Student Characteristics | | | | | | | | | |
| Asian/Pac. Isl. | 0.03 | 0.02 | 0.86 | 0.05 | 0.03 | 0.01 | 0.05 | 0.03 | 0.29 |
| Black | 0.19 | 0.20 | 0.03 | 0.11 | 0.13 | 0.00 | 0.13 | 0.10 | 0.77 |
| Latinx | 0.15 | 0.10 | 0.85 | 0.09 | 0.06 | 0.02 | 0.12 | 0.07 | 0.12 |
| white | 0.59 | 0.65 | 0.05 | 0.71 | 0.74 | 0.00 | 0.67 | 0.76 | 0.05 |
| Male | 0.52 | 0.47 | 0.41 | 0.39 | 0.38 | 0.79 | 0.50 | 0.51 | 0.97 |
| Age in 1994 (years) | 15.71 | 15.82 | 0.00 | 15.97 | 15.72 | 0.04 | 16.40 | 16.54 | 0.15 |
| Students with disabilities | 0.10 | 0.15 | 0.11 | 0.06 | 0.07 | 0.45 | 0.04 | 0.07 | 0.37 |
| English spoken at home | 0.92 | 0.94 | 0.89 | 0.94 | 0.96 | 0.04 | 0.92 | 0.98 | 0.98 |
| US Born | 0.92 | 0.96 | 0.99 | 0.93 | 0.96 | 0.07 | 0.92 | 0.97 | 0.91 |
| General health (0-4) | 0.84 | 0.80 | 0.31 | 0.86 | 0.80 | 0.01 | 0.89 | 0.81 | 0.03 |
| Log(Birthweight) | 4.51 | 4.53 | 0.62 | 4.77 | 4.76 | 0.03 | 4.77 | 4.78 | 0.69 |
| Times nominated as a best friend | 0.50 | 0.73 | 0.28 | 2.52 | 2.37 | 0.13 | 1.98 | 2.25 | 0.02 |
| Number of clubs and sports | 1.69 | 1.85 | 0.19 | 2.33 | 1.97 | 0.00 | 2.45 | 1.95 | 0.01 |
| Considered attractive, physically | 0.51 | 0.52 | 0.89 | 0.57 | 0.51 | 0.08 | 0.53 | 0.63 | 0.08 |
| Considered attractive, personality | 0.50 | 0.46 | 0.87 | 0.58 | 0.49 | 0.02 | 0.54 | 0.62 | 0.49 |
| Close relationship with mother | 0.87 | 0.85 | 0.18 | 0.85 | 0.86 | 0.91 | 0.83 | 0.82 | 0.60 |
| Always lived in same home | 0.28 | 0.28 | 0.11 | 0.25 | 0.26 | 0.56 | 0.34 | 0.17 | 0.02 |
| Biological father present | 0.88 | 0.90 | 0.66 | 0.94 | 0.89 | 0.00 | 0.92 | 0.89 | 0.52 |
| Panel B: Parent Characteristics | | | | | | | | | |
| Age in 1994 (years) | 43.01 | 42.04 | 0.16 | 42.28 | 41.18 | 0.86 | 41.73 | 41.43 | 0.96 |
| Disabled | 0.06 | 0.04 | 0.66 | 0.05 | 0.04 | 0.68 | 0.05 | 0.06 | 0.78 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| US Born | 0.83 | 0.90 | 0.20 | 0.88 | 0.92 | 0.01 | 0.92 | 0.94 | 0.03 |
| Recently accepted welfare | 0.08 | 0.09 | 0.29 | 0.05 | 0.07 | 0.43 | 0.08 | 0.03 | 0.24 |
| Neither parent has HS diploma | 0.11 | 0.17 | 0.02 | 0.09 | 0.10 | 0.06 | 0.09 | 0.05 | 0.54 |
| HS diploma highest deg. earned | 0.08 | 0.21 | 0.00 | 0.23 | 0.24 | 0.61 | 0.20 | 0.26 | 0.58 |
| Attended some college | 0.35 | 0.27 | 0.13 | 0.28 | 0.32 | 0.57 | 0.29 | 0.30 | 0.88 |
| Highest degree is bachelors | 0.21 | 0.24 | 0.95 | 0.18 | 0.19 | 0.73 | 0.28 | 0.22 | 0.12 |
| Graduate schooling | 0.25 | 0.11 | 0.00 | 0.21 | 0.15 | 0.04 | 0.13 | 0.16 | 0.65 |
| Household income in 1994 ($) | 63055 | 43588 | 0.00 | 54684 | 49629 | 0.03 | 48541 | 53471 | 0.70 |
| At least very good health | 0.59 | 0.46 | 0.00 | 0.56 | 0.55 | 0.03 | 0.54 | 0.53 | 0.14 |
| Smoker in household | 0.39 | 0.51 | 0.11 | 0.43 | 0.49 | 0.69 | 0.46 | 0.42 | 0.67 |
| Panel C: Census Tract Characteristics | | | | | | | | | |
| Population | 5529.51 | 5250.78 | 0.18 | 5375.81 | 5148.14 | 0.58 | 5476.82 | 4815.94 | 0.31 |
| Asian/Pac. Isl. | 0.06 | 0.02 | 0.06 | 0.04 | 0.03 | 0.00 | 0.03 | 0.02 | 0.00 |
| Black | 0.18 | 0.16 | 0.73 | 0.10 | 0.12 | 0.00 | 0.14 | 0.10 | 0.41 |
| Latinx | 0.12 | 0.06 | 0.17 | 0.08 | 0.04 | 0.02 | 0.11 | 0.05 | 0.69 |
| white | 0.71 | 0.79 | 0.50 | 0.82 | 0.83 | 0.00 | 0.79 | 0.85 | 0.29 |
| Pop. without HS diploma by 25 | 0.26 | 0.27 | 0.30 | 0.26 | 0.27 | 0.00 | 0.31 | 0.27 | 0.04 |
| Pop. without coll. degree by 25 | 0.26 | 0.22 | 0.03 | 0.23 | 0.22 | 0.73 | 0.20 | 0.23 | 0.03 |
| Household income ($) | 32792 | 29526 | 0.10 | 33257 | 30529 | 0.24 | 29775 | 30761 | 0.02 |
| Pop. on welfare | 0.09 | 0.09 | 0.89 | 0.08 | 0.08 | 0.00 | 0.10 | 0.08 | 0.02 |
| Owner occupied dwelling | 0.65 | 0.70 | 0.45 | 0.74 | 0.74 | 0.08 | 0.70 | 0.73 | 0.18 |
| Unemployment rate | 0.08 | 0.07 | 0.77 | 0.07 | 0.07 | 0.00 | 0.08 | 0.07 | 0.09 |
| n(students) | 202 | 1183 | | 340 | 1050 | | 112 | 444 | |

Note. Values represent portion of data unless otherwise noted. P-value compares students with a school-based mentor to those with no mentor. Add Health provided weights are used to achieve national representativeness.

## Appendix C

*Identifying romantic partner pairs.* We make iterative passes through romantic partner nominations by moving from the first to the third romantic partner nominations. The first step is we drop students that were missing data regarding natural mentorship. Next, we conduct our iterative process. Consider student A who nominates students 1, 2, and 3. In our process, we first consider student 1's first nomination. If this results in a matched pair between student A and student 1, then a unique romantic pair ID is assigned to the pair and both students are removed from the sample. Next, the iterative process would begin again with the remaining unreplaced, unmatched sample. If a pair is not created, we then consider if student A is student 1's second nomination and so on. We repeat this process for all of student A's nominated romantic partners in Wave I before repeating the process with student A's romantic partner nominations from Wave II. Whenever a requited pair is established, both students are removed from the sample before the next iteration of matching occurs. Thus, student A may be matched and removed from the sample before an alternative match involving student A could have occurred. There are a few instances where ties occur, in which case we randomly select a pair. Importantly, our sample of romantic pairs is not the unique solution for pairing requited romantic partnerships.

## Appendix D

We create our school-level estimates of students' sense of belonging using a standardized jackknife average of the following five item Likert response survey questions from Wave I:

H1ED18) How often do you have issues getting along with other students? {never; just a few times; about once a week; almost everyday; everyday}
H1ED19) You feel close to people at your school. {strongly agree; agree; neither agree nor disagree; disagree; strongly disagree}
H1ED20) You feel like you are part of your school. {same as above}
H1ED21) Students at your school are prejudiced. {same as above}
H1ED22) You are happy to be at your school. {same as above}
H1ED23) The teachers at your school treat students fairly. {same as above}
H1ED24) You feel safe at your school. {same as above}

We code all items such that higher values indicate a greater sense of belonging. Next, we conduct a principal components analysis across these 7 items and take the first component. We standardize this student-level measure of belonging to have a mean of zero and unit variance. We then estimate the leave-one-out school level averages as shown in model 1 below:

$$Belonging_{i,s} = \frac{1}{n_s - 1} \sum_{j=1,\, j \neq i}^{n_s} x_{j,s}$$

Where $Belonging_{i,s}$ is the jackknife estimate of belonging for student *i* in school *s* and $x_{j,s}$ is the first component of the PCA on the 7 belonging items mentioned above for student $j$ in school *s.* For students who had not yet entered high school by Wave I, we use the Wave I school-wide belonging average. To create school level estimates of belonging, we average the belonging measure above across all students in a school and then standardize these averages across schools. The resulting school level belonging has mean of 0 and variance of 1.