DISCUSSION PAPER SERIES

IZA DP No. 10319

# More Than Just Friends? School Peers and Adult Interracial Relationships

Luca Paolo Merlino
Max Friedrich Steinhardt
Liam Wren-Lewis

October 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

# More than Just Friends? School Peers and Adult Interracial Relationships

**Luca Paolo Merlino**
*Université Paris 1 Panthéon Sorbonne*

**Max Friedrich Steinhardt**
*Helmut Schmidt University,*
*IZA, LdA and CELSI*

**Liam Wren-Lewis**
*Paris School of Economics and INRA*

Discussion Paper No. 10319
October 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 10319
October 2016

# ABSTRACT

# More than Just Friends? School Peers and Adult Interracial Relationships*

This paper investigates the impact of individuals' school peers on their adult romantic relationships. In particular, we consider the effect of quasi-random variation in the share of black students within an individual's cohort on the percentage of adults' cohabiting partners that are black. We find that more black peers leads to more relationships with blacks later in life. The results are similar whether relationships begun near or far from school, suggesting that the racial mix of schools has an important and persistent impact on racial attitudes.

JEL Classification: J12, J15, J16

Keywords: romantic relationships, assortative matching, race

Corresponding author:

Liam Wren-Lewis
Paris School of Economics
48 Boulevard Jourdan
75014 Paris
France
E-mail: liam.wrenlewis@parisschoolofeconomics.eu

* We would like to thank Joseph Altonji, David Card, Scott Carrell, Pierre-André Chiappori, Margherita Comola, David Donaldson, Ben Elsner, Gabrielle Fack, Dan Hamermesh, Camille Hémet, Hilary Hoynes, Ingo Isphording, Rucker Johnson, Patrick Kline, Victor Lavy, Conrad Miller, Kaivan Munshi, Amanda Pallais, Thomas Piketty, Elena Stancanelli, Christopher Walters and participants at LAGV 2016, SOLE 2016, and seminars at Antwerp, Bicocca, IZA, Paris I, PSE and UC Berkeley for helpful comments and suggestions. The usual disclaimer applies.

# 1 Introduction

Interracial marriage and cohabitation rates are important indicators of social integration and the health of race relations (Fryer, 2007). In the US, where racial tensions are an ongoing problem,[1] it may therefore be concerning that interracial marriage between blacks and other races remains low. Hitsch, Hortaçsu and Ariely (2010), for instance, note that black men are over 10 times more likely to marry a black woman than they would under random matching within cities.

This paper investigates one potentially important factor explaining assortative matching by race: exposition to racial diversity at a young age. In particular, we explore how plausibly exogenous variation in a student's school peer group influences the romantic relationships they later undertake as an adult. Analyzing the impact of school cohorts allows us to draw important inferences on the deeper factors that explain variation in adult relationships. Moreover, school composition is one of the most direct ways that policy makers can influence racial integration.

Previous work on assortative matching by race has found an important part can be explained by racial preferences. For instance, Wong (2003) and Hitsch et al. (2010) use calibrated matching models to show that differences in endowments or meeting opportunities are unlikely to fully explain observed racial sorting. Fisman, Iyengar, Kamenica and Simonson (2008) furthermore show that race is important in determining the willingness to date in a speed dating experiment. Yet little is known about what determines these racial preferences, nor to what extent they are influenced by individuals' experiences.

Social interaction has long been postulated as a potential means of reducing racial prejudices (e.g. Williams, 1947; Allport, 1954). Indeed, recent studies have shown that white students exposed to a greater number of black students adjust their stated attitudes and choose to interact more frequently with blacks in schools (Marmaros and Sacerdote, 2006; Boisjoly, Duncan, Kremer, Levy and Eccles, 2006; Camargo, Stinebrickner and Stinebrickner, 2010; Carrell, Hoekstra and West, 2015).[2] It is somewhat unclear from existing papers, however, whether this effect extends beyond the particular contexts studied. Baker, Mayer and Puller (2011), for instance, show that variation in the racial composition of university dorms doesn't impact students' broader social networks. Moreover, these papers study the impact on limited interactions such

[1] See Moran and Waddington (2016) and The Economist (2016) for examples of recent manifestations of racial tensions.

[2] Additionally, Dobbie and Fryer (2015) show that white graduates who teach in schools where most students are black show increased racial tolerance.

as emailing or sharing a dorm, and hence it is yet to be demonstrated that such attitude changes affect economically important decisions such as cohabitation.

The main contribution of this paper is therefore to provide evidence that the racial composition of students' school cohorts impacts romantic relationships later in life. To do so, we use the National Longitudinal Survey of Adolescent Health (Add Health), which collects information on the race of all students within surveyed schools in the US and then over a decade later surveys a sample of these students on their romantic partners. This data allows us to exploit idiosyncratic variation in grade composition within schools, a methodology pioneered by Hoxby (2000) and widely used to identify causal peer effects.[3] A number of tests confirm that the variation we use is indeed random and uncorrelated with other variables that might influence adult relationships.

Our central result is that the share of adults' cohabiting partners who are black is increasing in the share of their school cohort who are black. The cohorts which impact adult relationships are students of the same sex in the same grade, and it is therefore not simply the result of students having more potential black partners in school. The magnitude of the effect is important—for whites, for instance, going from the average of 10 percent blacks in the cohort to 15 percent would imply increasing the black share of adult partners from the average of 2.8 percent to 3.6 percent. The result is robust to several robustness checks, including the introduction of grade-school fixed effects. Furthermore, we find no evidence that our results are driven by measurement error in the way outlined in Angrist (2014). We therefore conclude that school racial composition has an important impact on adult interracial relationships.

We then demonstrate that the most likely mechanism behind this impact is a change in racial preferences or attitudes. First, we find significant effects on reported attitudes in several waves of the survey. Second, consistent with friendships being the driver of attitude changes, we find a significant impact of cohorts on the racial composition of individuals' closest friends. Moreover, the effect of cohorts on adult relationships is stronger in schools which are less racially segregated, as defined by Echenique and Fryer (2007). Third, we document evidence suggesting that an increase in meeting opportunities through friends met in school is unlikely to play a major role. In particular, if our result stemmed mainly from increased meetings with blacks through school-based social networks, we would expect it to be stronger for those relationships formed in school, at a younger age and geographically closer to

[3]See, for instance, Bayer, Hjalmarsson and Pozen, 2009; Bifulco, Fletcher and Ross, 2011; Lavy, Paserman and Schlosser, 2012; Carrell, Hoekstra and Kuka, 2016, among others.

school. We find no evidence for such a differential impact. Fourth, we show that any impact of cohort racial composition on educational performance or labor market outcomes could not be large enough to explain our measured effect. Overall, therefore, our results suggest that racial diversity in schools impacts individuals' attitudes or beliefs which in turn affect their decisions regarding relationships.

We proceed in the following way. Section 2 details the dataset and estimation strategy, and provides evidence for the validity of our main identification assumption. In Section 3 we then present the benchmark results before proceeding to undertake a number of robustness checks, including adding additional controls, looking for bias driven by measurement error and considering alternative specifications. We then investigate our results further in Section 4 in an attempt to shed light on the mechanisms that may potentially be driving the result. Finally, Section 5 concludes and makes suggestions for future research.

# 2 Data and estimation strategy

## 2.1 Data

We use data from the National Longitudinal Survey of Adolescent Health (Add Health).[4] The survey selected 80 nationally representative high schools and 54 feeder schools in the US and first gave a questionnaire to all students in the school in grades 7-12 in 1994-95. This in-school survey was self-administered and collected basic information from around 90,000 students, including their gender and race. Within each school a sample of students was then interviewed at home and asked many detailed questions on topics including family background, health behaviors, friendships and romantic relationships. This in-home survey was administered to around 20,000 students and these students formed the sample for the following waves, administered in 1996 (Wave 2), 2001-02 (Wave 3) and 2008-09 (Wave 4).

In a first step, we use the in-school survey to construct information about

[4]The Add Health project was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (Add Health@unc.edu). No direct support was received from grant P01-HD31921 for this analysis.

school peers. In particular, we construct our main independent variable, the share of students in a peer group who are black.[5] We consider three alternative peer groups: the cohort of all students in the same grade, students of opposite sex in the same grade and the cohort of same sex students in the grade. A key advantage of using the in-school sample of the Add Health is that it is close to a census of students within the grade, and hence we reduce measurement error in cohort composition differences.

Our analysis then uses the Wave 4 in-home survey to measure outcomes in terms of relationships. Within this survey, respondents were asked to give basic information on a list of current and past romantic partners. This list included their current or most recent romantic partner as well as any person who they had been married to, had lived with for more than one month or had had a relationship with that resulted in pregnancy. We focus on relationships involving cohabitation, including marriage, since these are the most consistently recorded and represent the most committed relationships. With this information, we construct our main dependent variable: the share of an individual's reported cohabiting partners that are black.

Of the original Wave 1 sample, around 14,000 were interviewed and assigned sample weights in Wave 4.[6] Of this sample, we were unable to match 534 with information on their school cohort, and we dropped a further 101 for whom we observe less than five students in the in-school survey of the same gender.[7] Around 2,000 of the remaining individuals do not report any previous cohabiting partners in Wave 4, and therefore our main analysis uses a sample of around 11,000 students who report having cohabited with at least one partner.[8]

In terms of attrition, work by Harris (2013) suggests that attrition bias in Wave 4 is small in magnitude for demographic, behavioral, health and

[5]In the in-school survey, race is self-reported and students could define themselves as being of more than one race. In the analysis that follows the black share is defined as the share of students who defined themselves to be black. We consider alternative definitions of race in the robustness checks (Appendix C). For romantic partners listed in Wave 4, individuals can only report one race.

[6]The in-home survey sampled students with unequal probability, and hence we use sampling weights in our analysis. For more details on the Add Health data, see Chen and Chantala (2014).

[7]This is done in order to reduce noise stemming from the extreme values of our independent variable that these observations produce. Results are robust to the inclusion of these observations.

[8]We find no evidence that our dependent variables of interest impact the probability of having a cohabiting partner. As a robustness check, we also consider an alternative specification where our dependent variable is whether an individual has ever cohabited with a partner of a different race.

attitudinal variables after study estimates were adjusted with final sampling weights. Moreover, Bifulco et al. (2011) find no evidence that attrition is correlated with minority shares within cohorts. Indeed, when we regress a dummy for whether an observation is in our final sample on the grade black share, controlling for school fixed effects, we obtain a t-statistic of -1.16. We therefore conclude that attrition bias is unlikely to affect our results.

Summary statistics of the main variables we use in our analysis are reported in Table 1, along with other characteristics that help to describe our sample. We report the estimated population mean of a range of variables along with the estimated population standard deviations both between and within schools. Detailed variable definitions are given in Appendix A.

The relative scarcity of interracial relationships is immediately apparent: for the average white individual in our sample, around 3 percent of reported partners are black, whilst for black individuals this figure is over 80 percent. This disparity is more pronounced than in the school cohorts, where white students' cohorts contain an average of 10 percent blacks, and black students' cohorts contain an average of 53 percent blacks. The average within-school standard deviation in the grade black share is around 2 percentage points. If we restrict individuals' cohorts to be only those students of the same gender, this standard deviation increases to around 3 percentage points.

Individuals range between the ages of 11 and 21 when surveyed in Wave 1, with 29 percent attending a middle school and 57 percent attending a high school.[9] At Wave 4 individuals are aged between 24 and 34 and have cohabited with 1.4 partners on average.

## 2.2 Estimation strategy

We cannot simply regress dating behavior on cohort composition since cohort composition is likely to be correlated with a range of other omitted variables that impact dating behavior—not least, the composition of the population that lives nearby the school. Moreover, self-selection of individuals might further bias results if those who are more inclined to date blacks choose to go to schools with a larger share of black students.

In order to control for these omitted variables, we exploit variation in the

[9]We define a middle school as one which contains no grades higher than grade nine, and a high school as one which contains no grades lower than grade nine. Among those schools that contain both grades eight and ten, we find three schools with abnormally large increases between two grades (above 100 percent). In the analysis that follows, we follow a conservative approach, splitting each of these schools in two. Our results are however robust to not splitting these schools.

Table 1: Summary statistics

| | Mean | Within school s.d. | Between school s.d. | N |
|---|---|---|---|---|
| *Main variables* | | | | |
| Grade black share | .17 | .016 | .24 | 11353 |
| - for white subsample | .098 | .016 | .21 | 7579 |
| - for black subsample | .53 | .017 | .25 | 2357 |
| Share of cohabiting partners black | .15 | .24 | .23 | 11353 |
| - for white subsample | .027 | .12 | .04 | 7579 |
| - for black subsample | .82 | .28 | .34 | 2357 |
| *Other Wave 1 variables* | | | | |
| Age | 16 | 1.1 | 1.3 | 11353 |
| Female | .49 | .49 | .077 | 11353 |
| Race = White | .75 | .29 | .28 | 11353 |
| Race = Black | .16 | .22 | .25 | 11353 |
| Race = Asian | .032 | .1 | .09 | 11353 |
| Hispanic | .12 | .21 | .19 | 11353 |
| Family income | 46 | 35 | 21 | 8571 |
| Grade size | 216 | 28 | 136 | 11353 |
| Grades in school | 3.7 | 0 | 1.3 | 11353 |
| In middle school | .29 | 0 | .46 | 11353 |
| In high school | .57 | 0 | .5 | 11353 |
| Lives in urban area | .52 | .17 | .43 | 11248 |
| Region = Northeast | .16 | 0 | .37 | 11353 |
| Region = Midwest | .32 | 0 | .46 | 11353 |
| Region = South | .39 | 0 | .49 | 11353 |
| Region = West | .13 | 0 | .35 | 11353 |
| *Other Wave 4 variables* | | | | |
| Age | 29 | 1.2 | 1.3 | 11353 |
| Number of recorded partners | 2 | 1.4 | .32 | 11353 |
| Number of cohabiting partners | 1.4 | 1.1 | .27 | 11353 |
| Number of marriages | .55 | .55 | .19 | 11353 |
| Attended college | .65 | .45 | .16 | 11352 |
| Employed | .78 | .4 | .079 | 9434 |

*Notes:* Summary statistics are calculated using the Wave 4 cross-sectional weights, which aim to produce a representative sample from individuals who are surveyed in both Waves 1 and 4.

share of black students across cohorts within an individual school.[10] In other words, we assume that families do not select schools based on the differences between the average school composition and their child's school specific cohort and that these differences are not correlated with other imported omitted variables.

To implement our identification strategy, we estimate the following regression equation:

$$ShareBlackRelat_i = \alpha ShareBlack_{cs} + \beta X_i + I_{gm} + I_{sm} + \epsilon_i \qquad (1)$$

where $ShareBlackRelat_i$ is the share of individual $i$'s reported cohabiting partners that are black, $ShareBlack_{cs}$ is the share of blacks within cohort $c$ in school $s$, $X_i$ are a set of individual race dummies, $I_{gm}$ are grade-gender fixed effects, $I_{sm}$ are school-gender fixed effects, and $\epsilon_i$ is a random error term. We split school and grade fixed effects by gender since much of our analysis uses gender-specific cohort shares, and we are concerned of systematic differences in cohort shares across gender at the school and grade level. Standard errors are clustered at the school level. The main set of individual level controls include a Hispanic dummy and dummies for race, where race can be white, black, mixed, or other.

We run our benchmark regression on individuals of all race, including those who are black. The logic behind this is that, in general, the contact hypothesis does not focus specifically on the attitudes of the majority group (Sigelman and Welch, 1993). Indeed, interracial contact can affect the racial attitudes of both the majority and minority races. In our case, non-black students exposed to more black peers may change their racial attitudes and end up dating more blacks, and equally black students exposed to more non-black peers may change their racial attitudes and end up dating fewer blacks. While in the benchmark regression we keep individuals of all races, we will then run the regression separately for blacks and whites to investigate whether the effects differ across the two groups.

In our first regression, we consider an individual's peer group as the cohort of students in the same grade within the school. We then split grades into two, first considering students in the same grade of the opposite gender, and then considering the cohort of students in the same grade of the same gender. On the one hand, we may expect opposite sex peers to have the largest impact

[10]Schools with no black students are then going to be left out from our analysis. To examine the characteristics of the schools that contribute most to our results, Table D7 in Appendix D provides summary statistics for schools that have within-school variation in the black cohort share above median.

on adult relationships, since this group forms a pool of potential romantic partners. On the other hand, same sex peers may be more important if this is the group from which close friends are most likely to be drawn.

## 2.3 Identification assumption

Our methodology relies on the assumption that variation in cohort composition within schools is essentially random, once we control for grade-gender fixed effects. We can test two important implications of this assumption. First, we can test whether within-school variation in the share of black students is correlated with predetermined individual level variables—a type of balancing test. Second, we can test for non-random clustering of black students across grades within schools.

To perform the balancing tests, we regress a range of predetermined student characteristics on the black share of their peer group, while controlling for school-gender, grade-gender and race fixed effects. Each characteristic is regressed separately on the black share of students in the whole grade, as well as the black share of students of opposite and same sex in each grade. The corresponding results, presented in Table 2, support our main identification assumption - none of the pre-determined variables are significantly correlated with any measure of the peer group black share.

A second way to test for random assignment is to check for non-random clustering in black shares across cohorts. The simple intuition is that, if variation is random, then an individual's race should be uncorrelated with that of their peers. As noted by Guryan, Kroft and Notowidigdo (2009), however, one cannot test for this by simply regressing an individual's race on that of their peers, because each individual is present in many others peer groups but necessarily not their own. We therefore undertake a number of tests designed to avoid this problem, including those proposed by Guryan et al. (2009), Stevenson (2015) and Caeyers and Fafchamps (2016). Details of these tests and results can be found in Appendix B. Overall, none of the tests reject random clustering, and we therefore conclude that the distribution of blacks after controlling for fixed effects is consistent with quasi-random variation.

Finally, the variation in share of black students across cohorts may be partly affected by the end of court-ordered desegregation orders which occurred during this time. Lutz (2011) show that the expiration of court oversight led to significant changes in racial composition, but these changes are not correlated with other trends, and hence this is not a threat to our identification. Moreover, Table 2 shows that there is no significant correlation between our variation and neighborhood black shares, suggesting that our variation is not

Table 2: Balancing tests for cohort composition measures

| | N | Independent variable: Grade black share, both genders | Grade black share, opp. gender | Grade black share, same gender |
|---|---|---|---|---|
| Age at Wave 4 | 13,849 | -0.455 | -0.181 | -0.317 |
| | | (0.387) | (0.261) | (0.291) |
| Parent is black | 11,906 | 0.0217 | 0.00650 | 0.0118 |
| | | (0.0559) | (0.0396) | (0.0405) |
| Share of census block black | 13,724 | 0.116 | 0.0655 | 0.0729 |
| | | (0.0855) | (0.0659) | (0.0689) |
| Share of neighbors black | 13,534 | -0.0368 | 0.00691 | -0.0407 |
| | | (0.105) | (0.0683) | (0.0592) |
| Gender ratio in grade | 13,849 | -0.103 | -0.156 | 0.00681 |
| | | (0.0666) | (0.146) | (0.169) |
| Grade size | 13,849 | 115.3 | 54.13 | 57.08 |
| | | (84.70) | (41.64) | (40.59) |
| Born in USA | 13,849 | 0.0107 | 0.0387 | 0.00816 |
| | | (0.102) | (0.0758) | (0.0874) |
| Lives with both biological parents | 11,885 | 0.0202 | -0.0576 | 0.0765 |
| | | (0.243) | (0.174) | (0.192) |
| Number of older siblings | 13,832 | -0.423 | -0.427 | 0.0417 |
| | | (0.672) | (0.422) | (0.424) |
| Years of parental schooling | 13,121 | 1.099 | 0.350 | 0.594 |
| | | (0.864) | (0.683) | (0.725) |
| Log of family income | 10,379 | 0.0248 | 0.255 | -0.263 |
| | | (0.423) | (0.357) | (0.299) |
| Home language is not English | 13,849 | -0.0668 | -0.0607 | 0.00298 |
| | | (0.0661) | (0.0595) | (0.0670) |

*Notes:* Each coefficient is from a seperate regression where the variable in the first column is regressed on one of the three specified independent variables, with controls including race, grade-gender fixed effects, and school-gender fixed effects. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3: benchmark Results

| | (1) | (2) | (3) |
|---|---|---|---|
| Grade black share, both genders | 0.197** | | |
| | (0.0970) | | |
| Grade black share, opposite gender | | -0.0672 | |
| | | (0.0823) | |
| Grade black share, same gender | | | 0.276*** |
| | | | (0.0643) |
| Observations | 11353 | 11318 | 11353 |
| Adjusted $R^2$ | 0.715 | 0.715 | 0.716 |

*Notes:* The table reports OLS estimates. The dependent variable is share of cohabiting partners who are black. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

being driven by changes in the areas students are taken from or changes in the racial composition of those areas.

# 3 Main results

Results of the benchmark regressions are presented in Table 3. The coefficient on grade black share in column 1 is significant at the 5 percent level, providing evidence that individuals who had a greater share of blacks in their cohort on average have more black partners later in life.

Columns 2 and 3 present the results when we split grades by gender. We can see that the coefficient on the black share of the opposite sex peer group is insignificantly different from zero, whereas the coefficient on the black share of the same sex cohort is positive and highly significant. The result in column 1 is therefore not being driven by students meeting romantic partners in their own grade. This is perhaps unsurprising given that a student's grade is less likely to be the relevant population for within-school romantic relationships than for friendship. Indeed, in our data, a majority of students' within-school friends are in the same grade, but this is true for less than a quarter of within-school romantic partners.

Since it is the composition of same sex students which is driving our re-

sults, from now on we focus on these cohorts. The importance of same sex peers could be driven by the fact that students typically form closer friendships with students of their own gender, and hence this peer group has the largest impact on post high-school relationships.[11] Before we explore in detail potential mechanisms for this relationship, we first explore the robustness of this result.

We first analyze the extent to which our results may be affected by omitted variable bias by introducing a series of additional control variables. These results are presented in Table 4. Column 1 presents our same sex cohort benchmark result, with column 2 introducing a number of control variables, including grade size and the number of partners an individual has reported cohabiting with. Column 3 then additionally adds a set of control variables which reduce the sample size, including the share of the census block which is black and log family income. In each case, our coefficient of interest remains almost unchanged, suggesting that our result is not being driven by unobservables correlated with the controls we add (Altonji, Elder and Taber, 2005; Oster, 2016).

We can additionally control for a number of unobservables by introducing school trends and other fixed effects. Column 4 adds a trend variable for each school, which partially controls for factors such as differential dropout rates among blacks across schools that are not picked up by grade-gender fixed effects. Column 5 adds interaction terms between school fixed effects and race dummies, which controls for differential responses to school context between blacks and whites. Column 6 then introduces census tract fixed effects along with their interaction with race dummies. There are an average of 25 census tracts within a school, so including census tract fixed effects further ensure our results are not being driven by variation in the residential area from which students are drawn. Finally, column 7 adds in school-grade fixed effects interacted with race dummies, such that our coefficient of interest is identified entirely from the difference between the black share among students of the same gender and the black share among students of the opposite gender. In all these regressions, our coefficient moves relatively little while remaining highly significant. Overall, estimates in Table 4 strongly suggest that unobserved omitted variables are unlikely to be driving our result.

Next, we explore to what extent our results vary across different subsamples within our population. At first, we split our sample by race. The corresponding

[11] In line with this, Soetevent and Kooreman (2007) find that interactions with peers of the same gender are generally much stronger than those with peer of the opposite gender for several academic and non-academic outcomes.

Table 4: Robustness to additional controls

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Grade black share, same gender | 0.276*** (0.0643) | 0.277*** (0.0654) | 0.270*** (0.0962) | 0.226*** (0.0825) | 0.255*** (0.0642) | 0.354*** (0.0944) | 0.332*** (0.112) |
| Extended controls 1 | | Y | Y | | | | |
| Extended controls 2 | | | Y | | | | |
| School trends | | | | Y | | | |
| School-race FE | | | | | Y | | |
| Tract-race FE | | | | | | Y | |
| School-grade-race FE | | | | | | | Y |
| Observations | 11353 | 11353 | 7845 | 11353 | 11353 | 11254 | 11353 |
| Adjusted $R^2$ | 0.716 | 0.717 | 0.723 | 0.718 | 0.734 | 0.777 | 0.752 |

*Notes:* The table reports OLS estimates. The dependent variable is share of cohabiting partners who are black. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Extended controls in column 2 include language spoken at home, grade size, cohort size, number of reported cohabiting partners, whether an individual lived with a single parent at Wave 1, whether an individual was born in the US and the individual's age at Wave 4. In column 3 we further include controls for the census block population black share, parental race, years of parental education, whether an individual lives with both biological parents at Wave 1, the number of older siblings and log family income. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

estimates in columns 1 and 2 of Table 5 show that our main result holds for both whites and blacks. Comparing the two coefficients using a Wald test we cannot reject the hypothesis that the coefficient is the same in the two regressions.

Columns 1 and 2 also allow us to get a sense of the importance of the magnitude of the impact. The point estimates imply that, for whites, going from the average of 10 percent blacks in the cohort to 15 percent would increase the average black share of adult partners from 2.8 percent to 3.6 percent. For blacks, moving from 58 percent blacks in the cohort to the sample average of 53 percent would reduce the average black share of adult partners from 82 percent to 80 percent.[12]

[12]Ideally we would also like to compare the magnitude of the effect to the impact of some other variable, but we are not aware of any factor influencing interracial relationships which has been well-identified. Instead, in column 3 of Appendix Table C4, we additionally include the black share of an individual's census block in Wave 1, and find the coefficient to be significant but about half the size of the coefficient on the cohort black share.

Table 5: Subsample splits

| | Race | | School type | | School segregation | |
|---|---|---|---|---|---|---|
| Sample: | Whites | Blacks | Middle | High | $\leq$ median | $>$ median |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Grade black share, same gender | 0.157* | 0.387** | 0.489** | 0.235*** | 0.594*** | 0.202*** |
| | (0.0875) | (0.189) | (0.210) | (0.0844) | (0.170) | (0.0648) |
| P-val, coefs equal | .31 | | .25 | | .03 | |
| Observations | 7579 | 2357 | 2374 | 6780 | 5970 | 5325 |
| Adjusted $R^2$ | 0.077 | 0.334 | 0.699 | 0.719 | 0.314 | 0.749 |

*Notes:* The table reports OLS estimates. The dependent variable is the share of cohabiting partners who are black. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

In columns 3 and 4 of Table 5 we split the sample between middle and high schools. We note that the coefficients on cohort black share are significant in both cases, providing further evidence that our results are not driven by differential dropout rates or students moving between schools. The two coefficients are not significantly different, and if anything the impact of middle school cohorts appears slightly stronger. This is consistent with an impact on attitudes, where we can imagine exposition at a younger age has a higher impact, and suggests the result is not being driven by students meeting future partners in school or through friendships formed in school.

The final split shown in Table 5 is by the extent to which blacks are segregated within the school. This is based on friendship networks of blacks within the school and is constructed according to the methodology of Echenique and Fryer (2007). When we divide the sample in this way, we find that the impact of an individual's cohort black share is significantly higher within schools that are less segregated. One potential explanation for this is that diversity within a students cohort has most impact when students are more likely to befriend students of another race, and this is less likely in more segregated schools. It should be noted, however, that school segregation as well as school black share are correlated with many other variables, and therefore we should not over-interpret this difference.

A range of other subsample splits are displayed in Appendix C. In general we find splitting the sample in other ways, such as by student gender or school black share, does not yield notable differences in results.

One potential concern with the methodology of exploiting cohort variation is that results can be driven by selection bias or measurement error, as described by Angrist (2014).[13] In our setting, the concern would be that our measures of race do not perfectly measure an individual's 'true' race, and this true race is correlated with the measured race of their peers. Since we have strong evidence that our variation in cohort black shares is random, this is unlikely to be a problem in our case, but we also check for bias from measurement error in three ways. First, if the cohort black share was proxying for an individuals' true race, we would expect it to be significant when we replace our dependent variable with predetermined variables correlated with race. Yet in Table 2, we can see that the coefficient is insignificant when regressing parental race and two alternative measures of neighborhood black shares. Second, if our result is driven by measurement error in race, the coefficient should fall when we introduce other variables correlated with race, yet we observe little change when we introduce a dummy for whether the interviewer in Wave 4 declares the individual to be black. Third, following Feld and Zölitz (2016) and Carrell et al. (2016), we redo our estimation introducing varying amounts of measurement error. As expected, a greater amount of measurement error leads to results being attenuated to zero, and does not bias the coefficient upward. Results from these tests and more details can be found in Appendix C.

A different concern is that, since our identification is driven by small random variation across cohorts, our standard errors may be inappropriate or our results may be driven by some other aspect of the cohort which is correlated with the black share. We can test for this in two ways. First, we construct over two hundred other cohort shares including, for instance, the share of Hispanics and the share who have college educated mothers. We first include these all simultaneously as control variables and observe that there is little impact on our coefficient of interest. We then enter them into regressions individually in place of our main explanatory variable and record the t-statistic. The t-statistic with the largest magnitude in this resulting distribution is -2.62, while the t-statistic of our coefficient of interest in the benchmark regression with the same gendered cohort is 4.30 (see Figure C2 in Appendix C). Second, we undertake one thousand 'placebo' regressions whereby we assign students to cohorts within their school at random. Plotting the distribution of coefficients, we note that the true coefficient is clearly an outlier as it is larger than any of

[13] Angrist (2014) also discusses a range of other potential problems with the peer effects literature, but these are mainly not relevant to our context. For instance, non-linearity is not an issue since our peer characteristic of interest—being black or not—is binary in nature.

the placebo coefficients (see Figure C3 in Appendix C). We can therefore conclude that it is very unlikely that our result is driven by chance or correlation with another characteristic of school cohorts.

Finally, we may be concerned that our results are sensitive to the particular specification that we have chosen. We use the share of partners that are black as our dependent variable as this is a simple specification that allows us to treat black and non-black individuals symmetrically. This specification, however, implicitly assumes that the cohort black share has no impact on the probability of cohabiting. We find no evidence that the cohort black share has any such impact, but as an alternative we consider a specification where the dependent variable an indicator for whether an individual has ever cohabited with someone of a different race, where race is defined as black or non-black. In this regression, the coefficient on the cohort black share is still positive and significant, and indeed our result is robust to a number of alternative specifications, including looking at other types of relationships and using alternative definitions of being black. We also check whether we find similar effects when considering other races, but we find no such evidence when we consider minorities other than blacks. Details of these results can be found in Appendix C.

Overall, we have so far established that students who have more black pupils in their school cohort on average go on to have more relationships with blacks as adults. This result is consistent with Gordon and Reber (2016), who find evidence that school racial desegregation between 1961 and 1985 led to more black-white births. Their results are sensitive to the specification of cohort trends, however, and our work is therefore complementary in providing evidence based on an alternative source of variation. In the next section, we exploit additional data available in the Add Health to provide insight on the potential mechanisms behind our results.

# 4 Investigating mechanisms

There are three mechanisms that could lie behind the result identified in the previous section. First, the effect may be the result of a change in individuals' attitudes. Greater exposure to blacks - or less exposure to non-blacks - may change beliefs in line with the 'contact hypothesis'. Second, the effect may be the result of increased meeting opportunities. A greater number of blacks within an individual's cohort may increase the number of blacks within their social network, and through this network they may meet a greater number of potential black partners. Third, the effect may be the result of poorer

educational achievement. Various studies suggest that an increased share of black students may worsen educational achievement for their peers (Hoxby, 2000; Hanushek, Kain and Rivkin, 2009; Billings, Deming and Rockoff, 2014), and this may have knock-on effects on college attendance or employment which in turn impact their propensity to cohabit with blacks.

Identifying the mechanism at work is important for understanding the nature of peer effects as well as for policy. In particular, if adult relationships change as a result of a change in attitudes, this suggests cohort composition may impact a broader ranger of behavior including discrimination. In this section, we aim to investigate whether evidence from our data is consistent with one or more of these mechanisms.

## 4.1 Attitudes

To test whether our result is compatible with the contact hypothesis, we would like to look directly for changes in attitudes regarding race. The Add Health surveys do not ask questions specifically about such attitudes, but in Wave 3 respondents are asked to rate how important they think several elements are for a serious committed relationship. One of these elements is "being of the same race or ethnic group". We construct a binary measure of the relative importance of race in a relationship by comparing the rating given to race to the other factors. Since the question asks about being of the same race, we split our sample by respondents' race. The coefficient is significant and negative for white respondents, indicating that whites who had a greater share of blacks in their school cohort attach less importance to racial homogeneity within romantic relationships.

We would also like to get a sense as to whether attitudes towards race are impacted beyond the context of romantic relationships, and for this we employ two strategies. First, in the Wave 1 in-school survey students are asked how much they agree with the general statement "the students at this school are prejudiced". Answers to this question could incorporate any form of prejudice, but for black respondents we can imagine answers should partly reflect the extent to which they feel non-black students are prejudiced towards black students. Column 3 of Table 6 reports the results of regressing this variable on the cohort black share, and we indeed find a greater share of blacks within a grade leads black respondents to report less prejudice.

A second strategy involves exploiting the fact that in Waves 3 and 4 individuals are asked how politically liberal they are. In general changes in racial attitudes are unlikely to significantly shift individuals' overall political identification, and indeed when we regress individuals' political identification

Table 6: Impact of cohort black share on attitudes

| Dependent variable: | Importance of race in relationships | | Are students prejudiced? | Change in liberalness in Obama election year |
|---|---|---|---|---|
| Sample: | Whites | Blacks | Blacks | Whites |
| | (1) | (2) | (3) | (4) |
| Grade black share, same gender | -0.713*** | -0.168 | -0.688*** | 1.484** |
| | (0.253) | (0.262) | (0.260) | (0.626) |
| Observations | 6300 | 1887 | 12125 | 5477 |
| Adjusted $R^2$ | 0.074 | 0.055 | 0.083 | 0.040 |

*Notes:* The table reports OLS estimates. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Wave 4 cross-sectional weights are used in all regressions. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

in Wave 3 on cohort black share the coefficient is insignificant. Wave 4 of the survey, however, was undertaken in 2008, the year when Barack Obama was first campaigning to become president. In both the democratic primary and the general election, Obama positioned himself as the more liberal of the candidates with large-scale support, and we may therefore hypothesize that individuals' political identification in 2008 may be particularly correlated with their attitudes towards blacks.[14] In column 4, we therefore regress the change in declared liberalness between Waves 3 and 4 on the cohort black share. The coefficient is positive and significant, suggesting that a greater share of blacks among white students' peers impacts their general attitudes towards blacks.

[14]To test the hypothesis that change in declared liberalness over this period is correlated with attitudes towards blacks, we analyzed panel data covering the 2008 election provided by the American National Election Studies. Surveys in this panel asked both questions on racial attitudes and a question on political identification identical to that asked in Add Health. Consistent with our hypothesis, we find changes in respondents' political identification during the period were positively correlated with the degree of sympathy they felt to blacks. Furthermore, for non-blacks in our Add Health sample, the change in liberalness variable is significantly correlated with the importance given to race within relationships and the share of black friends, even when we control for cohort black share. More details are available on request.

## 4.2 Friendships

Both the contact hypothesis and the meeting opportunities hypothesis suggest that there should be an impact on friendships in addition to romantic relationships. Since Add Health collects information on friendships in Waves 1 and 2, we can directly test for an impact of cohort black share on the share of school friendships that are with black students. Moreover, in Wave 4 the survey asks whether individuals' close friends are of the same race as them. To test for an impact of cohort composition on friendships, we regress these measures on the share of blacks amongst students of the same gender and grade. The results are displayed in Table 7.

Table 7: Impact of cohort black share on friendships

| Dependent variable: | Share of Waves 1 & 2 friends black | Share of Waves 1 & 2 reciprocal friends black | All Wave 4 close friends are of same race as individual | |
|---|---|---|---|---|
| Sample: | All | All | Whites | Blacks |
| | (1) | (2) | (3) | (4) |
| Grade black share, same gender | 0.177* | 0.410*** | 0.0604 | 0.790** |
| | (0.102) | (0.137) | (0.251) | (0.323) |
| Observations | 8675 | 5389 | 7394 | 2259 |
| Adjusted $R^2$ | 0.744 | 0.711 | 0.101 | 0.160 |

*Notes:* The table reports OLS estimates. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Column 1 shows that there is a significant effect of the racial composition of same sex cohort on the racial composition of nominated friends in Waves 1 and 2. Column 2 considers only friends who also nominated the individual as a friend—Lavy and Sand (2012) finds that these 'reciprocal' friends have stronger effects on their peers than other friends. Again, the effect is significant and positive.

When we consider close friends in Wave 4 in columns 4 and 5, the black share of the same gender cohort has a significant impact for blacks. In particular, blacks whose school grade had a higher proportion of blacks among students of the same gender are more likely to only have black close friends as adults. There is no significant effect on non-blacks, but this is perhaps not surprising given that the question is about friends being of the same race, without

particular reference to blacks. Overall, therefore, the evidence on friendships is consistent both with the contact hypothesis and the meeting opportunities mechanisms.

## 4.3 Impact by distance from school

Our results so far are consistent with a change in racial attitudes, but it is also possible that the effect we have found is driven by an increase in meeting opportunities that result from social networks formed in school. Indeed, we have seen that students with more black peers make more black friends and hence it might be that, through these friends, they then meet more other black people who become romantic partners. In order to test whether such meeting opportunities can explain most of the impact on adult cohabitation, in this section we investigate heterogeneity in our result by different measures of distance from school. The basic idea is that, if the result is driven by social networks formed in school, then the effect we have identified should be strongest for those relationships closest to school. On the other hand, if greater cohabitation with blacks is driven by a change in attitudes, then our results should be similar when we consider relationships that began far from school.

We can consider three different ways in which relationships can be far from school. First, for relationships undertaken whilst students are in school, we can consider whether or not their partner attended the same school. Second, relationships can be distant from school in time—partners met a long time after an individual has left school are less likely to have been met through school friends. Third, adult relationships can be distant geographically—relationships begun in different states are less likely to have been formed through school-based social networks. For each of these types of distance, therefore, we would expect the 'meeting opportunities' channel to be relatively weak for those relationships begun far away.

Table 8 reports the results of our standard regression when we split the sample of partners in three different ways. In columns 1 and 2, we consider relationships reported in Waves 1 and 2, and divide partners according to whether or not they attended the same school as the respondent. The dependent variable is the share of each set of partners who are black. Comparing the two coefficients, we see that there is no evidence that the effect of cohort black share is strongest for partners within school. Indeed the effect is highly significant for relationships with partners outside of school, with a substantially

Table 8: Impact by distance from school

| Relationship characteristic: | Relationships reported in Waves 1 & 2 | | Marriages and cohabitations | | | |
|---|---|---|---|---|---|---|
| | Partner in same school | Partner not in same school | Begun before age 22 | Begun after age 22 | Begun in school state | Begun out of school state |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Grade black share, same gender | 0.0883 | 0.337** | 0.163 | 0.436*** | 0.256*** | 0.175 |
| | (0.152) | (0.151) | (0.120) | (0.116) | (0.0690) | (0.201) |
| P-val, coefs equal | .13 | | .14 | | .71 | |
| Observations | 6245 | 5438 | 6521 | 7574 | 9403 | 2805 |
| Adjusted $R^2$ | 0.733 | 0.748 | 0.688 | 0.685 | 0.714 | 0.683 |

*Notes:* The table reports OLS estimates. The dependent variable is the share of (different types of) relationships that were with black people. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

larger coefficient.[15]

In columns 3 to 6 we return to cohabitations reported in Wave 4 as our relationships of interest. Columns 3 and 4 consider whether our effect differs in temporal distance from school by separating relationships according to whether they began before or after the agent was aged 22. If the effect is being driven by meeting opportunities stemming from school social networks, it would likely be most important in relationships formed before or just after students left school. This is not what we find, however, with the point estimate being substantially larger for those relationships formed longer after leaving.

Finally, to analyze the differential impact by geographical distance we would ideally like to know where partners met. Unfortunately, the data does not contain such information, so instead we exploit information provided by respondents on when, if ever, they moved between US states. We combine this with information on when the relationship began to form a set of relationships that began after the respondent left the state they went to school in. In column 6, the dependent variable is then the share of these relationships that are with blacks, with column 5 considering the black share of the remaining relationships. Though the coefficient on the cohort black share is insignificant in column 6, it is not significantly different from that in column 5. Overall,

[15]We can similarly split the same relationships by whether the partner and respondent have a mutual friend. Doing so produces similar results, with a larger coefficient in the regression restricted to partners who do not have a mutual friend with the respondent.

Table 9: Impact of cohort black share on educational performance

| | Average test score (1) | Attended college (2) | Employed (3) | Earnings (4) |
|---|---|---|---|---|
| Grade black share, same gender | -0.199 | 0.0375 | -0.164 | -3938.4 |
| | (0.282) | (0.164) | (0.154) | (12096.1) |
| Observations | 11243 | 11352 | 9434 | 11150 |
| Adjusted $R^2$ | 0.129 | 0.135 | 0.076 | 0.109 |

*Notes:* The table reports OLS estimates. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

therefore, we can conclude that there is little evidence to suggest a differential impact by distance, and hence the meeting opportunities mechanism alone is unlikely to be driving our results.

## 4.4 Educational performance

It is reasonable to hypothesize that a student's performance in school may have an impact on the race of their future adult partners. For instance, if worse grades mean students are less likely to go to college, they may then meet proportionally more black people, and as a result be more likely to cohabit with blacks.

To test for this mechanism, we first analyze whether we observe any impact of black cohort share on average test scores, college attendance, employment or earnings. The results of these regressions are presented in Table 9, and we can see that the coefficient on the black share is always insignificant. This is consistent with Bifulco et al. (2011) who find no impact of minority shares on these outcomes.

The insignificance of the results in Table 9 may however result from a lack of power rather than the absence of any real effects. Indeed, an impact on test scores of the size estimated in Billings et al. (2014), for instance, is within our 95 percent confidence interval. We therefore regress our main outcome variable on these measures of education, employment and earnings to come up with an approximate upper bound for the size of this mechanism.[16]

[16]Results available on request. One reason this upper bound is approximate is because the

Even if we assume that all of the true coefficients are at the upper bounds of the various 95 percent confidence intervals, we estimate that impacts on educational performance, employment and earnings can account for no more than 9 percent of the effect identified in our benchmark. We therefore conclude that educational performance is unlikely to be an important mechanism in explaining our result.

# 5 Conclusions

This paper finds that greater racial diversity in schools significantly impacts the prevalence of interracial adult relationships. The higher the share of black classmates of the same gender, the more likely an adult has a black partner during adulthood. Moreover, we provide suggestive evidence that most of this effect is driven by changes in attitudes. This indicates that policies designed to increase racial diversity in schools, such as busing programs, may be at least somewhat effective in reducing racial prejudices and encouraging social integration more generally. Indeed, they may also help to reduce assortative matching on race in the marriage market, and hence potentially reduce household income inequality.

Our findings suggest a number of interesting avenues for future work. One question which we are unable to answer is to what extent the middle or high-school contexts are particularly important for influencing adult relationships. On the one hand, adolescence may be a crucial period since it is the time when individuals are beginning to explore romantic relationships, and afterwards racial attitudes may be harder to change. On the other hand, diverse social contexts may be more important earlier in life, when racial attitudes are presumably weaker, or later in life, when individuals form long-term romantic relationships.

Our paper has also highlighted the need to investigate further whether being exposed to racial diversity at school has implications for racial attitudes outside of the social sphere. One could imagine that attitudes that impact friendship networks and romantic relationships might also affect discrimination in the labor market or the workplace, but it is also possible that the latter may be more impacted by work-related experiences. This is an important question to investigate if we are to understand fully the impacts of policies designed to increase racial diversity.

---

coefficients in this regression are likely to be biased due to a number of omitted variables. We nonetheless use these coefficients because no well-identified impacts of these variables exist, and the direction of this bias is in any case likely to inflate the result.

# References


Allport, Gordon W, (1954), *The Nature of Prejudice*, Reading, MA : Addison-Wesley Pub. Co.

Altonji, Joseph G, Todd E Elder, and Christopher R Taber, (2005), "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, *113* (1), 151–184.

Angrist, Joshua D, (2014), "The Perils of Peer Effects," *Labour Economics*, *30*, 98–108.

Baker, Sara, Adalbert Mayer, and Steven L Puller, (2011), "Do More Diverse Environments Increase the Diversity of Subsequent Interaction? Evidence from Random Dorm Assignment," *Economics Letters*, *110* (2), 110–112.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen, (2009), "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections," *The Quarterly Journal of Economics*, *124* (1), 105–147.

Bifulco, Robert, Jason M Fletcher, and Stephen L Ross, (2011), "The Effect of Classmate Characteristics on Post-secondary Outcomes: Evidence from the Add Health," *American Economic Journal: Economic Policy*, *3* (1), 25–53.

Billings, Stephen B., David J. Deming, and Jonah Rockoff, (2014), "School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg," *The Quarterly Journal of Economics*, *129* (1), 435–476.

Boisjoly, Johanne, Greg J Duncan, Michael Kremer, Dan M Levy, and Jacque Eccles, (2006), "Empathy or Antipathy? The Impact of Diversity," *The American Economic Review*, *96* (5), 1890–1905.

Caeyers, Bet and Marcel Fafchamps, (2016), "Exclusion Bias in the Estimation of Peer Effects," NBER Working Paper.

Camargo, Braz, Ralph Stinebrickner, and Todd Stinebrickner, (2010), "Interracial Friendships in College," *Journal of Labor Economics*, *28* (4), 861–892.

Carrell, Scott E, Mark Hoekstra, and Elira Kuka, (2016), "The Long-Run Effects of Disruptive Peers," NBER Working Paper 22042.

_ , _ , and James E West, (2015), "The Impact of Intergroup Contact on Racial Attitudes and Revealed Preferences," NBER Working Paper 20940.

Chen, Ping and Kim Chantala, (2014), "Guidelines for Analyzing Add Health Data," *Carolina Population Center, University of North Carolina at Chapel Hill.*

Dobbie, Will and Roland G Fryer, (2015), "The Impact of Voluntary Youth Service on Future Outcomes: Evidence from Teach For America," *The BE Journal of Economic Analysis & Policy*, *15* (3), 1031–1065.

Echenique, Federico and Roland G Fryer, (2007), "A Measure of Segregation Based on Social Interactions," *The Quarterly Journal of Economics*, *122(2)*, 441–485.

Feld, Jan and Ulf Zölitz, (2016), "Understanding Peef Effects - On the Nature, Estimation and Channels of Peer Effects," *Journal of Labor Economics*, *forthcoming.*

Fisman, Raymond, Sheena S Iyengar, Emir Kamenica, and Itamar Simonson, (2008), "Racial Preferences in Dating," *The Review of Economic Studies*, *75* (1), 117–132.

Fryer, Roland G, (2007), "Guess Who's Been Coming to Dinner? Trends in Interracial Marriage over the 20th century," *The Journal of Economic Perspectives*, *21* (2), 71–90.

_ , Lisa Kahn, Steven D Levitt, and Jörg L Spenkuch, (2012), "The Plight of Mixed-Race Adolescents," *Review of Economics and Statistics*, *94* (3), 621–634.

Gordon, Nora and Sarah Reber, (2016), "The Effects of School Desegregation on Mixed-Race Births," NBER Working Paper 22480.

Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo, (2009), "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," *American Economic Journal: Applied Economics*, *1* (4), 34–68.

Hanushek, Eric A, John F Kain, and Steven G Rivkin, (2009), "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement," *Journal of Labor Economics*, *27* (3).

Harris, Kathleen Mullan, (2013), "The Add Health Study: Design and Accomplishments," Technical Report, Carolina Population Center University of North Carolina at Chapel Hill.

Hill, Andrew J, (2015), "The Girl Next Door: The Effect of Opposite Gender Friends on High School Achievement," *American Economic Journal: Applied Economics*, *7* (3), 147–177.

Hitsch, Günter J, Ali Hortaçsu, and Dan Ariely, (2010), "Matching and Sorting in Online Dating," *The American Economic Review*, *100(1)*, 130–163.

Hoxby, Caroline, (2000), "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER Working Paper 7867.

Lavy, Victor and Edith Sand, (2012), "The Friends Factor: How Students' Social Networks Affect Their Academic Achievement and Well-Being?," NBER Working Paper 18430.

_ , M Daniele Paserman, and Analia Schlosser, (2012), "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom," *The Economic Journal*, *122* (559), 208–237.

Lutz, Byron, (2011), "The End of Court-Ordered Desegregation," *American Economic Journal: Economic Policy*, *3(2)*, 130–168.

Marmaros, David and Bruce Sacerdote, (2006), "How Do Friendships Form?," *The Quarterly Journal of Economics*, *121* (1), 79–119.

Moran, Matthew and David Waddington, (2016), "Back to the Future: Race and Riots in Ferguson, Missouri," in Matthew Moran and David Waddington, eds., *Riots: An International Comparison*, London: Palgrave Macmillan UK, pp. 141–170.

Oster, Emily, (2016), "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, *http://dx.doi.org/10.1080/07350015.2016.1227711.*

Sigelman, Lee and Susan Welch, (1993), "The Contact Hypothesis Revisited: Black-White Interaction and Positive Racial Attitudes," *Social Forces*, *71* (3), 781–795.

Soetevent, Adriaan R and Peter Kooreman, (2007), "A Discrete-Choice Model with Social Interactions: with an Application to High School Teen Behavior," *Journal of Applied Econometrics*, *22* (3), 599–624.

Stevenson, Megan, (2015), "Breaking Bad: Mechanisms of Social Influence and the Path to Criminality in Juvenile Jails," *Available at SSRN 2627394.*

The Economist, (2016), "Race in America: Progress and its discontents," *July 16th.*

Williams, Robin M., (1947), "The Reduction of Intergroup Tensions: a Survey of Research on Problems of Ethnic, Racial, and Religious Group Relations.," *Social Science Research Council Bulletin.*

Wong, Linda Y, (2003), "Why do only 5.5% of Black Men Marry White Women?," *International Economic Review*, *44* (3), 803–826.

# Appendix A Variable definitions

Table A1: Description of Variables

| Variable | Wave | Description | Values |
|---|---|---|---|
| *Main variables* | | | |
| Grade black share | 1 | Share of students in an individual's grade who define themselves to be black | [0,1] |
| Share of cohabiting partners black | 4 | Share of an individual's reported cohabiting partners that are black | [0,1] |
| *Other variables* | | | |
| Are students prejudiced? | 1 | Extent to which students agree with statement "The students at this school are prejudiced" | 0,...,4 |
| Average test score | 1 | Average of most recent grade in maths, english, history and science | [1,4] |
| Change in liberalness in Obama election year | 3 & 4 | Answer to question "In terms of politics, do you consider yourself very conservative, conservative, middle of-the-road, liberal, or very liberal?" in Wave 3 subtracted from answer to same question in Wave 4 | -4,...,4 |
| Earnings | 4 | Income received from personal earnings before tax | in U.S.$ |
| Family income | 1 | Annual family income of individual | in thous. U.S.$ |
| Gender ratio of grade | 1 | Number of students of opposite gender to individual in individual's grade, over number of students of same gender | [0,3.2] |
| Grade size | 1 | Number of students in individual's grade | 9,...,695 |
| Grades in school | 1 | Number of grades in individual's school | 1,...,6 |
| Importance of race in relationships | 3 | Takes value 0 if 'being of the same race or ethnic group' is ranked as a less important element of a serious relationship than love, fidelity, commitment and money; takes value 1 otherwise | 0,1 |
| In middle school | 1 | Individual's school contains no grades beyond grade 9 | No = 0, Yes = 1 |
| In high school | 1 | Individual's school contains no grades before grade 9 | No = 0, Yes = 1 |
| Living in urban area | 1 | Respondent lives in an urban area | No = 0, Yes = 1 |
| Number of cohabiting partners | 4 | Number of partners individual cohabited with for at least a month | 0,...,21 |
| Parent is black | 1 | Parent interviewed in Wave 1 (normally resident mother) defines themselves as black | No = 0, Yes = 1 |
| School segregation | 1 | Segregation of blacks in individual's school, as defined by Echenique and Fryer (2007) | [0,1] |
| Share of census block black | 1 | Proportion of census block population black | [0,1] |
| Share of neighbors black | 1 | Weighted proportion of the 20 students in the Add Health in-home survey who live closest that are black, calculate using algorithm created by Hill (2015) | [0,1] |
| Years of parental schooling | 1 | Number of years in school of individual's most educated parent | 8,...,17 |

# Appendix B Tests for non-random clustering

We undertake a number of tests that look for evidence of non-random clustering of black students within schools in the in-school survey data. The intuitive idea behind these tests is that, if cohorts are more or less black in some systematic way, then an individual's race will be significantly correlated with that of their peers. Traditionally, this hypothesis would be tested by regressing a dummy variable of whether an individual is black on the black share of the rest of their cohort. However, it is now understood that such a test will typically produce a negatively biased coefficient since an individual's peer group necessarily does not contain the individual themselves (Guryan et al., 2009; Angrist, 2014; Stevenson, 2015; Caeyers and Fafchamps, 2016). We therefore undertake several tests that have been designed to avoid this 'exclusion bias'.

Caeyers and Fafchamps (2016) derive analytically a formula for the exclusion bias and then show that one can test for non-random clustering by transforming the standard test appropriately. In particular, in column 1 of Table B2 we use as a dependent variable the 'transformed black dummy' $\widehat{Black_i}$, where

$$\widehat{Black_i} = Black_i - bias_{cs} \times ShareBlack_{cs} \tag{2}$$

Here $Black_i$ is a dummy taking the value 1 if individual $i$ is black, and $bias_{cs} = \frac{(N_s-1)(K_c-1)}{(N_s-1)(N_s-K_c)+(K_c-1)}$, where $N_s$ is the number of students in the school and $K_c$ the number of students in the cohort. The regression produces a small negative insignificant coefficient, and hence cannot reject random clustering. In columns 2 and 3, we then carry out a similar test for the smaller peer groups—namely students in the same grade of opposite and same gender respectively. For peers of opposite gender, there is no need to transform the dependent variable since peers are picked from a different set of students to the individual. In both of these columns, coefficients are also small and insignificant, consistent with our assumption of quasi-random allocation across grades.

An alternative method for correcting for exclusion bias is proposed by Guryan et al. (2009), who suggest controlling for the set of all potential peers. In our case, this involves adding the black share among all other students in the school as a control variable. Results of this test are displayed in columns 4 and 5, and again we note that the coefficients of interest, on the cohort black shares, are insignificant.

A simple less formal test is presented in column 6, whereby we regress the male black share on the female black share. Again, the coefficient is insignificant, suggesting that there is unlikely to be important self-selection or omitted variables when it comes to race shares, since most factors which we

Table B2: Tests for non-random clustering

| | Transformed black dummy (1) | Black dummy (2) | Transformed black dummy (3) | Black dummy (4) | Black dummy (5) | Black share of males in grade (6) |
|---|---|---|---|---|---|---|
| Black share of others in grade | -0.0524 | | | 0.144 | | |
| | (0.197) | | | (0.284) | | |
| Black share of opposite gender in grade | | -0.00388 | | | | |
| | | (0.0725) | | | | |
| Black share of others of same gender in grade | | | 0.0566 | | -0.0971 | |
| | | | (0.125) | | (0.158) | |
| Black share of others in school | | | | -72.49*** | -72.25*** | |
| | | | | (19.02) | (19.01) | |
| Black share of females in grade | | | | | | -0.00286 |
| | | | | | | (0.0671) |
| Observations | 81897 | 80954 | 81897 | 81897 | 81897 | 80955 |
| Adjusted $R^2$ | 0.979 | 0.422 | 0.929 | 0.516 | 0.516 | 0.986 |

*Notes:* The table reports OLS estimates. Controls in columns 1 to 5 include grade-gender fixed effects and school-gender fixed effects, and in column 6 include grade and school fixed effects. Regressions reported in this table are run on the Wave 1 in-school survey. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

could imagine influencing the female black share would also simultaneously influence the male black share.

Stevenson (2015) suggests an alternative test for non-random clustering, which involves randomly picking one observation within each cohort and regressing the share of blacks among the rest of the cohort on a dummy for whether the selected individuals are black, along with school-gender and grade-gender fixed effects. In this way, each observation is only present on either the RHS or LHS in each regression and there is no bias generated. We do this 10,000 times and, using her derived test statistics, obtain a p-value of .53 for the grade black share and a p-value of .48 for the same gender grade black share. This suggests we are far from being able to reject random clustering.

Finally, Feld and Zölitz (2016) show that if variation in cohort black shares is systematic, then measurement error will bias the results of our regressions upwards. On the other hand, if the variation is random, then measurement error will bias our results downwards. In Section C we test for the impact of measurement error on our results by introducing random error in our measure of race, and we find that doing so biases our results towards zero. This is

therefore most consistent with variation in cohort black shares being random.

# Appendix C Supplementary results and robustness checks

## Additional subsample splits

In Table C3 we present further subsample splits that are of interest to understand further any variation in the impact of black students on adult romantic relationships. First, columns 1 and 2 show that the effect is significant for both male and female students, and not significantly different. We also find no evidence of significant differences when we break down the sample by race and gender simultaneously in columns 3 to 6. Moreover, there is no significant difference when we split the sample by mother's education in columns 7 and 8. Since these splits involve splitting the sample within schools, they substantially reduce our power, and therefore our ability to detect significant differences between subsamples is relatively small.

The second half of the table splits the sample by school characteristics. From columns 9 and 10 we can note that the result is similar for schools having more or less than 15 percent blacks, respectively, while columns 11 and 12 demonstrate the result holds for both relatively large and small grades. Finally, columns 13 to 16 demonstrate that there are significant differences across region, with a larger coefficient in the West and a very small coefficient in the Northeast. This may be related to the segregation results found in Table 5, since segregation is higher in the South and North-East than the West and Midwest.

## Robustness to measurement error

As discussed in Section 3, if the variation in black shares within schools is not random, then the estimated coefficients might be biased due to measurement error. One way to test whether our results may be biased by measurement error, suggested by Feld and Zölitz (2016) and Carrell et al. (2016), is to gradually introduce measurement error into our data and observe how our coefficient of interest changes. In particular, we repeat the following process 1000 times. First, we generate a new variable which takes the value one with a probability equal to the predicted black share based on school, gender and grade. Second, we generate a new black dummy variable which takes the observed value with a 99 percent chance and the random value with a 1 percent

Table C3: Additional subsample splits

*Split by individual characteristics*

| | Gender | | Gender & race | | | | Mother college educated | |
|---|---|---|---|---|---|---|---|---|
| Sample: | Female (1) | Male (2) | White females (3) | White males (4) | Black females (5) | Black males (6) | Yes (7) | No (8) |
| Grade black share, same gender | 0.200* (0.106) | 0.336*** (0.117) | 0.177 (0.141) | 0.121 (0.0765) | 0.317* (0.188) | 0.486 (0.353) | 0.320*** (0.111) | 0.179 (0.112) |
| P-val, coefs equal | .44 | | .63 | | | | .44 | |
| Observations | 6205 | 5148 | 4091 | 3488 | 1358 | 999 | 5911 | 5442 |
| Adjusted $R^2$ | 0.715 | 0.721 | 0.076 | 0.040 | 0.346 | 0.268 | 0.721 | 0.720 |

*Split by school characteristics*

| | School black share | | Grade size | | Region | | | |
|---|---|---|---|---|---|---|---|---|
| Sample: | $< 15\%$ (9) | $> 15\%$ (10) | $< 200$ (11) | $> 200$ (12) | North-east (13) | Mid-west (14) | South (15) | West (16) |
| Grade black share, same gender | 0.307** (0.138) | 0.276*** (0.0756) | 0.291*** (0.0798) | 0.286** (0.115) | 0.0653 (0.276) | 0.390*** (0.125) | 0.215*** (0.0697) | 0.893*** (0.203) |
| P-val, coefs equal | .84 | | .97 | | .01 | | | |
| Observations | 6912 | 4441 | 6060 | 5293 | 2581 | 2877 | 4426 | 1469 |
| Adjusted $R^2$ | 0.344 | 0.730 | 0.696 | 0.741 | 0.438 | 0.695 | 0.786 | 0.462 |

*Notes:* The table reports OLS estimates. The dependent variable is the share of cohabiting partners who are black. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Wave 4 cross-sectional weights are used Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

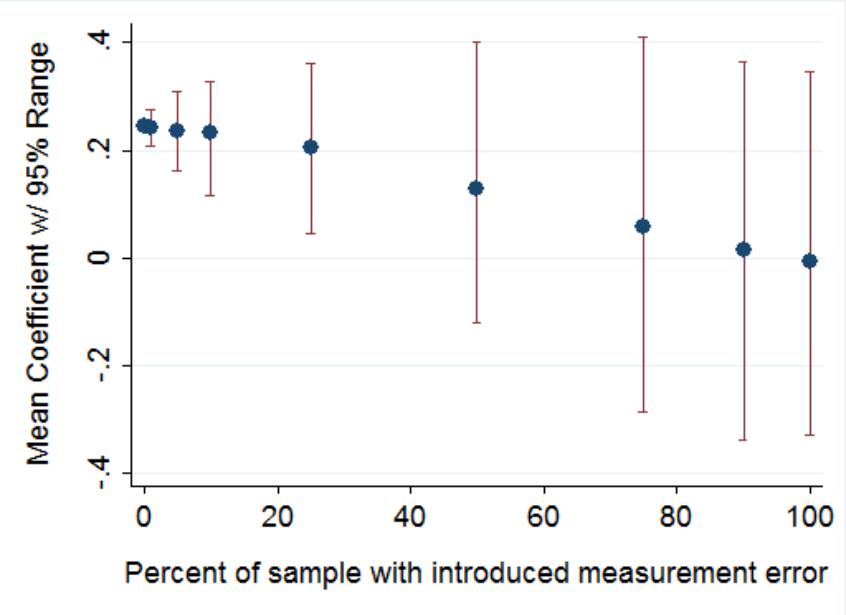
* Significant at the 10 percent level.

chance. We construct new cohort black shares based on this dummy and then undertake our standard regression with race variables besides the black dummy removed. Third, we repeat this process for other error levels.

Figure C1 shows the results of this process, where we plot the average coefficient generated as well as the 95 percent range. We can see that as more measurement error is introduced, the coefficient falls towards zero. This is consistent with our variation being random, and shows that measurement error would bias our results downwards rather than upwards.

Another way to check for measurement error is to add variables that may be correlated with the measurement error and observe whether our result changes. We therefore add to our benchmark regression two variables that are likely to be correlated with an individual's true race: a dummy for whether the interviewer in Wave 4 declares the surveyed individual to be black, and the share of the population that are black in the census block where they live. The result are shown in columns 2 and 3 of Table C4, and we include our

Figure C1: Sensitivity of main result to measurement error in black variable




*Notes:* The y-axis variable is the average coefficient on cohort black share from 1000 regressions where, before each regression, the black dummy variable is replaced with a random value for a share of the sample. This share is indicated on the x-axis.

benchmark regression in column 1 for comparision. Both added variables are positive and highly significant, but the coefficient on the cohort black share changes little from the benchmark result in column 1. This further suggests that measurement error is unlikely to be driving our results.

One further technique that has been used to address the concerns of Angrist (2014) is to split the sample between the individuals who may be producing the peer effects from those who are being influenced by them. In column 3, therefore, we show that our result holds when we restrict to the white individuals and make the main independent variable the number of blacks in their cohort.

## Placebo tests

To address concerns that our standard errors may be inappropriate or our results may be driven by some other cohort characteristic, we undertook two different sets of placebo tests. First, we constructed over two hundred other cohort share variables based on other questions in the in-school survey. The

Table C4: Robustness to measurement error and addition of other grade shares

| | All (1) | All (2) | Whites (3) | All (4) | All (5) |
|---|---|---|---|---|---|
| Grade black share, same gender | 0.276*** | 0.266*** | 0.273*** | | 0.272*** |
| | (0.0643) | (0.0672) | (0.0661) | | (0.0678) |
| Black dummy - Wave 4 interviewer | | 0.263*** | | | |
| | | (0.0535) | | | |
| Census block black share | | | 0.124*** | | |
| | | | (0.0272) | | |
| Number of blacks of same gender | | | | 0.000656* | |
| | | | | (0.000372) | |
| Other same gender grade shares | | | | | Y |
| Observations | 11353 | 11353 | 11244 | 7579 | 11353 |
| Adjusted $R^2$ | 0.716 | 0.721 | 0.718 | 0.076 | 0.718 |

*Notes:* The table reports OLS estimates. The dependent variable is share of cohabiting partners who are black. Controls include race, grade-gender fixed effects, and school-gender fixed effects. Column 4 additionally includes other measures of cohort composition derived from the in-school survey. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

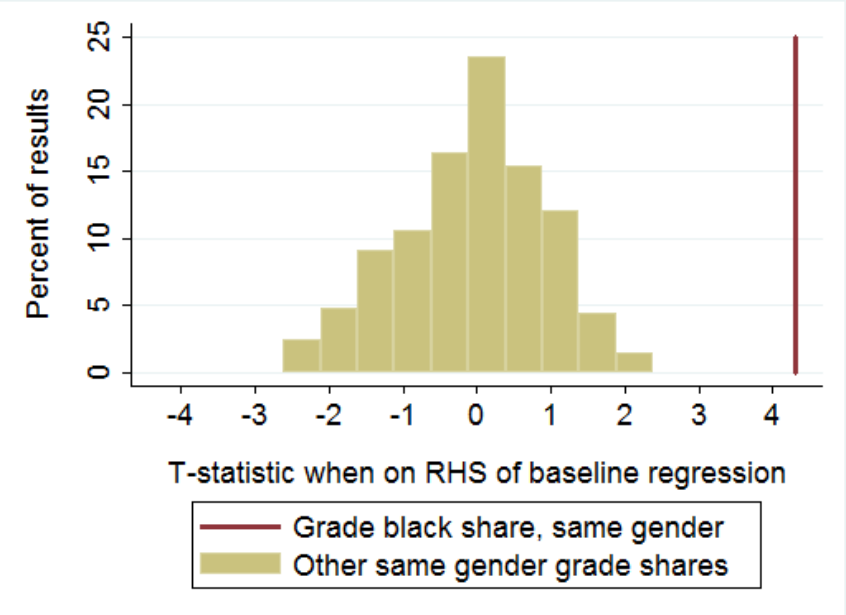
** Significant at the 5 percent level.

* Significant at the 10 percent level.

resulting variables include, for instance, the share of the cohort who are hispanic, the share who live with both of their parents, and the share whose most recent history grade was an A. Figure C2 then plots the t-statistics from the regressions when we enter each of these variables individually into our regression instead of the cohort black share. The red line represents the t-statistic we obtain in our benchmark, and this is clearly an outlier. Moreover, we add all of these variables to our benchmark. The results are displayed in column 5 of Table C4 and we can note that it has little impact on our main result. Hence, we can conclude that it is very unlikely that our result is driven by chance or correlation with another characteristic of school cohorts.

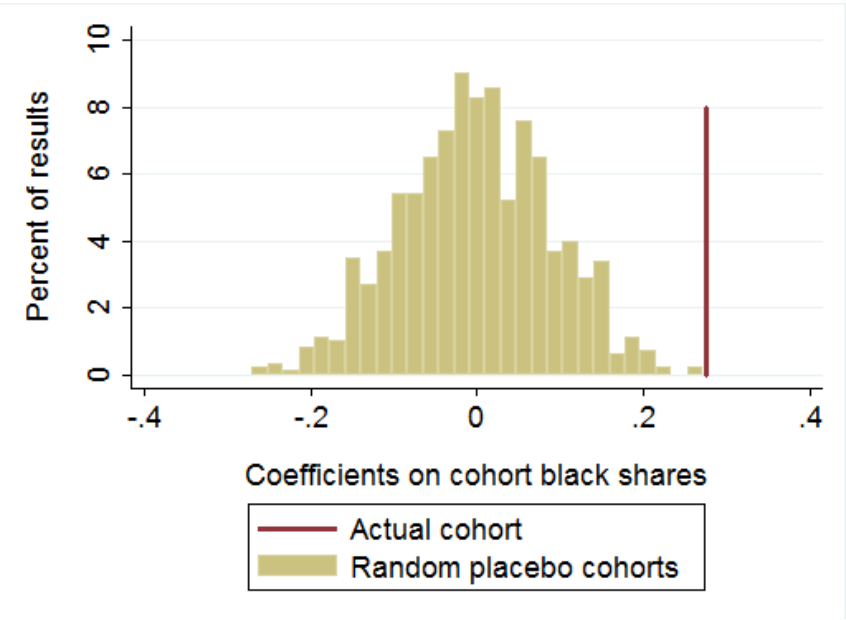
Our second placebo test involves reassigning students to cohorts randomly so that our measure of cohort black share is in general not that of their cohort, but another within the same school. We then carry out the same regression as in our benchmark one thousand times to produce a distribution of coefficients,

FIGURE C2: DISTRIBUTION OF T-STATISTICS FROM REGRESSIONS ON OTHER COHORT SHARES




*Notes:* Each t-statistic is taken from a regression where the dependent variable is the share of cohabiting partners who are black, and the independent variable is one of over 200 cohort share variables. The red line is the t-statistic in the benchmark, i.e. column 3 of Table 3.

Figure C3: Distribution of coefficients from regressions on randomly assigned cohort shares




*Notes:* Each coefficient is taken from a regression where the dependent variable is the share of cohabiting partners who are black, and the independent variable is the share of blacks among a randomly chosen cohort within the school. The red line is the coefficient in the benchmark, i.e. column 3 of Table 3.

which is displayed in Figure C3 alongside the coefficient from our benchmark. We can note from the figure that the distribution is centered on zero, as expected, and of the thousand placebo regressions none produces a coefficient as large as that from our benchmark. This further confirms that our result is not spurious.

## Other relationship measures

Table C5 presents the results of making a number of alternative relationship measures the dependent variable in our standard specification. Column 1 shows that the impact of the cohort black share has a very similar effect on partners as recorded in Waves 1 and 2, further confirming that our main result is unlikely to be driven by changes in college attendance or employment. We use Wave 1 weights in this regression so as not to restrict the sample to those re-interviewed in Wave 4, and this thus brings additional evidence that our

Table C5: Other relationship measures

| Dependent variable: | Share of partners black, Waves 1 & 2 (1) | Ever cohabited (2) | Number of cohabitations (3) | Ever married (4) | Share of non-cohabiting partners black (5) | Share of children with blacks (6) |
|---|---|---|---|---|---|---|
| Grade black share, same gender | 0.219** | -0.164 | 0.0469 | -0.0286 | 0.197 | 0.236** |
| | (0.0841) | (0.141) | (0.406) | (0.188) | (0.153) | (0.108) |
| Observations | 13518 | 13518 | 13518 | 13518 | 5587 | 6728 |
| Adjusted $R^2$ | 0.773 | 0.071 | 0.059 | 0.144 | 0.705 | 0.730 |

*Notes:* Controls include race, grade-gender fixed effects, and school-gender fixed effects. Wave 1 cross-sectional weights are used in column 1, all other columns use Wave 4 cross-sectional weights. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

baseline result is not being driven by attrition.

In columns 2 to 4 we use the large sample of all those who answered the relationship questions in Wave 4, rather than just those who reported having cohabited at least once. Column 2 shows that there is no significant impact on the probability of an individual having cohabited at least once, suggesting we do not face selection issues when focusing on the share of cohabitation. Column 3 shows that there is no significant impact on the number of cohabitions an individual has, while column 4 then shows there is no significant effect on whether an individual gets married. These results imply that the shift in the racial composition of relationships does not go along with any change in their nature. Column 5 examines the effect on the black share non-cohabiting partners, which we have previously excluded from the analysis since only a non-random selection of such partners are recorded. For this smaller sample, we find the coefficient to be of similar magnitude to our benchmark. Finally, in Column 6, we see that our result also holds when we examine the share of individuals' children that they have with black partners, which is consistent with the results of Gordon and Reber (2016).

## Other specifications of race

Table C6 presents the results when we look at race in different ways. First, in column 1, we use the sample of relationships where the partner was also interviewed—in Waves 1 & 2 a number of in-school partners were interviewed as part of the normal in-home survey, and in Wave 3 a random subsample of

current partners were interviewed. Using this subsample of partners allows us to use a different definition of partner race from that declared by the main interviewee, and thus examine the possibility that cohort black share influences interviewee's reporting of partners' race. Column 1 then shows that cohort black share still has a significant impact when we use the interviewer's report of the partner's race, and we obtain a very similar result when we use partners' self-reported race. Impacts on interviewees reporting of partners' race is therefore unlikely to be affecting our results.

Table C6: Other specifications of race

| Dependent variable: | Share of partners: | | | | | | Any partner of different race (Two races) (7) |
|---|---|---|---|---|---|---|---|
| | Black (per the interviewer) (1) | Black (2) | Hispanic (3) | Asian (4) | Of different race (Two races) (5) | Of different race (Four races) (6) | |
| Grade black share, same gender | 0.402** | | | | | | |
| | (0.175) | | | | | | |
| Grade black share (alt. def.), same gender | | 0.306*** | | | | | |
| | | (0.0728) | | | | | |
| Grade Hispanic share, same gender | | | -0.0807 | | | | |
| | | | (0.0907) | | | | |
| Grade Asian share, same gender | | | | -0.132 | | | |
| | | | | (0.158) | | | |
| Grade different race share, same gender | | | | | 0.427*** | | 0.388** |
| | | | | | (0.119) | | (0.174) |
| Grade different race share, same gender | | | | | | 0.0402 | |
| | | | | | | (0.0912) | |
| School-race FE | | | | | Y | Y | Y |
| Observations | 3076 | 11353 | 11353 | 11353 | 11353 | 11353 | 13849 |
| Adjusted $R^2$ | 0.770 | 0.716 | 0.404 | 0.374 | 0.211 | 0.301 | 0.084 |

*Notes:* The table reports OLS estimates. Partners in column 1 are all those interviewed in Waves 1,2 and 3; all other columns concern cohabiting partners reported in wave 4. Controls include race, grade-gender fixed effects and school-gender fixed effects. Columns 5 to 7 include school-race fixed effects. Wave 4 cross-sectional weights are used. Standard errors (in brackets) are clustered at the school level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

In column 2, our black share is calculated based on those who declare themselves to be black *and only black*, rather than simply those who declare themselves to be black. This distinction might be important since those of mixed race are likely to behave differently from blacks (Fryer, Kahn, Levitt

and Spenkuch, 2012). The coefficient is slightly larger but very similar, which is not surprising given the relatively small number of mixed race individuals. Columns 3 and 4 then look at two other minorities, Hispanics and Asians, and find no significant relationship between their cohort share and subsequent adult relationships with this groups. One potential explanation is that prevailing attitudes towards interracial relationships with these groups are different.

Columns 5 to 7 then take a different approach to our main specification by distinguishing between people of the same race and those of a different race. We start off with the specification closest to our benchmark which is to classify people into two races—black and non-black. In order to effectively control for the school racial composition, we must now include school fixed effects interacted with each race category. From column 5, we can see that the coefficient is still significant with this alternative specification, in line with the fact that we previously found our result held for both blacks and whites.

In column 6, we use a four-way definition of race—white, black, Asian and other—and find no effect, further suggesting that there is something particular to the black/non-black split. One potential explanation is that within-school variation in other race measures derive from differential responses in how individuals answer the question but not in how they are perceived. Indeed, in the Wave 1 in-home survey, only 2 percent of those who identify as black are not identified as black by the interviewer, but over 30 percent of those who identify as "other" are categorized as white, black or Asian by the interviewer. In column 7, we return to the two-race specification and focus on the extensive margin—whether or not individuals had any relationship with a person of a different race. Again we find the cohort black share to have a significant impact. Since this regression does not restrict to individuals who cohabited with at least one partner, it shows that this restriction is not driving our main result.

# Appendix D Schools with most variation in cohort black share

Table D7 provides summary statistics for schools that have within-school variation in the black cohort share above median. Comparing with Table 1, we see that the main difference is the share of black students in the school. We also see that these schools are more likely to be located in the South and have a greater share of students living in urban areas, but there are still many schools in areas outside the South and with rural students that are in this half of the sample.

Table D7: Summary statistics for schools with cohort black share s.d. above median

| | Mean | Between school s.d. | Within school s.d. | N |
|---|---|---|---|---|
| *Main variables* | | | | |
| Grade black share | .3 | .26 | .027 | 6242 |
| - for white subsample | .19 | .24 | .027 | 3292 |
| - for black subsample | .55 | .26 | .026 | 1959 |
| Share of cohabiting partners black | .24 | .25 | .33 | 6242 |
| - for white subsample | .038 | .051 | .15 | 3292 |
| - for black subsample | .84 | .24 | .31 | 1959 |
| *Other Wave 1 variables* | | | | |
| Age | 16 | 1.3 | 1.2 | 6242 |
| Female | .49 | .063 | .5 | 6242 |
| Race = White | .62 | .28 | .38 | 6242 |
| Race = Black | .26 | .28 | .33 | 6242 |
| Race = Asian | .035 | .079 | .12 | 6242 |
| Hispanic | .16 | .24 | .24 | 6242 |
| Family income | 43 | 21 | 35 | 4511 |
| Grade size | 209 | 140 | 29 | 6242 |
| Grades in school | 3.8 | 1.2 | 0 | 6242 |
| In middle school | .25 | .45 | 0 | 6242 |
| In high school | .6 | .5 | 0 | 6242 |
| Lives in urban area | .62 | .43 | .15 | 6189 |
| Region = Northeast | .15 | .36 | 0 | 6242 |
| Region = Midwest | .23 | .39 | 0 | 6242 |
| Region = South | .48 | .51 | 0 | 6242 |
| Region = West | .15 | .37 | 0 | 6242 |
| *Other Wave 4 variables* | | | | |
| Age | 29 | 1.3 | 1.2 | 6242 |
| Number of recorded partners | 2.1 | .35 | 1.5 | 6242 |
| Number of cohabiting partners | 1.4 | .3 | 1.1 | 6242 |
| Number of marriages | .52 | .2 | .55 | 6242 |
| Attended college | .62 | .19 | .45 | 6242 |
| Employed | .79 | .071 | .41 | 5193 |

*Notes:* Summary statistics are calculated using the Wave 4 cross-sectional weights, which aim to produce a representative sample from individuals who are surveyed in both Waves 1 and 4.