Low-Income Female Students and the Reversal of the Black-White Gap in High School Graduation

Brian Clark

U.S. Department of Justice

Ying Shi

Syracuse University

This article shows that the traditional narrative of Black-White high school graduation gaps is inverted among economically disadvantaged female students. Two nationally representative surveys and statewide administrative data demonstrate that low-income White females graduate at rates 5 to 6 percentage points lower than Black peers despite having higher test scores. Greater rates of tobacco, alcohol, and illicit drug use among White females account for one third of the attainment disparity. Since the early onset of substance use among low-income White females predicts lower attainment, more research on the factors leading to risky behaviors and their correlates during early adolescence is warranted. Examining racial gaps in high school graduation at the intersection of gender and income categories can inform more tailored interventions.

Introduction

Researchers have long documented widespread disparities in educational and labor market outcomes by race. On average, African Americans have lower test scores, lower educational attainment, and worse outcomes in terms of wages and labor market participation (Altonji & Blank, 1999; Bowen et al., 2009; Clotfelter et al., 2009; Fryer & Levitt, 2006; Murnane et al., 2006; Reardon & Robinson, 2008). Existing evidence shows that the Black-White gap in high school graduation is approximately 8 to 9 percentage points for those born in the 1980s (Heckman & LaFontaine [HL], 2010; Murnane, 2013). Attainment disparities carry significant implications for future economic well-being (Tamborini et al., 2015); for example, high school dropouts fail to capture the substantial college wage premium in the labor market (Card, 1999; Grogger & Eide, 1995; Heckman et al., 2006). Closing high school graduation gaps across different racial and ethnic groups is therefore of concern to researchers and policymakers.

In this article, we uncover a result not in accordance with perceived wisdom on racial disparities: economically disadvantaged Black girls complete high school at *higher* rates than their disadvantaged White peers. Graduation rates for low-income White girls are 5 to 6 percentage points lower than poorer Black girls across both national longitudinal surveys and statewide administrative data. We do not observe this reversal in racial gaps across any other gender and income category, including low-income boys. This finding underscores the importance of examining racial gaps at more

granular sociodemographic categories to inform tailored interventions on attainment.

We examine the influence of three sets of measures for this graduation deficit among low-income White females: variation in family, neighborhood, and schooling contexts; differences in cognitive skills; and a range of students' behavioral and socioemotional attributes that proxy for noncognitive skills. Specific measures include grade retention, absences, tardies, delinquency, self-esteem, substance use, and expectations. We operationalize the analyses using the National Longitudinal Survey of Youth 1997 (NLSY97) and the National Longitudinal Study of Adolescent to Adult Health (Add Health), which provide detailed family background attributes, schooling variables, and student behaviors and attitudes. We supplement these national samples with administrative statewide data from the North Carolina Education Research Data Center (NCERDC) and its variables on student standardized test scores, class-level, and school characteristics.

Our results show that students' behavioral and socioemotional attributes matter for explaining the attainment deficit among low-income White girls. The use of alcohol, tobacco, and illicit drugs accounts for one third of the overall disparity. Low-income White girls exhibit markedly higher rates of substance use across both national surveys, with the largest differences in the use of illicit drugs. The concentration of these behaviors among the low-income White population is concerning given previous research on the sensitivity of non-Hispanic Whites' educational attainment to health-related

factors (Jackson, 2009). These behaviors not only have ramifications for educational outcomes (Chatterji, 2006; Marie & Zölitz, 2017) but they also have negative consequences for longer term health (Carpenter & Dobkin, 2009; Hall & Degenhardt, 2009). Our results point to concerning patterns among economically disadvantaged White populations in their adolescent, prelabor market years, which can meaningfully influence later life outcomes.

In contrast to the role of behavioral factors, analyses show that lower graduation rates among low-income White females remain after accounting for detailed family earnings, neighborhood and school characteristics, urban versus rural contexts, and cognitive skills. Income differences cannot explain this phenomenon, as both national surveys show that low-income White females come from higher earning households relative to Black peers. Moreover, lower graduation rates among low-income White females persist despite their higher standardized test scores and related measures of academic performance, such that accounting for cognitive skills exacerbate rather than narrow differences in attainment. These findings motivate a reexamination of the traditional narrative on racial attainment gaps among economically disadvantaged populations, since White students fail to translate advantages in these domains to higher attainment.

In examining an understudied phenomenon, this article contributes to a substantial literature on high school completion and dropout. Previous studies document high school graduation patterns using household surveys, including the Current Population Survey and the American Community Survey, the Common Core of Data, and longitudinal data sets. It is well established using these sources that Black students graduate at lower rates, with only a subset of studies documenting graduation and dropout for finer-grained subgroups such as by gender and race (Heckman & LaFontaine, 2010; Mishel & Roy, 2006; Murnane, 2013; Stillwell et al., 2011). Studies show that Black students reliably trail White students across high school graduation and completion rates. Similarly, event dropout rates show subgroups such as Black or low-income students as leaving school at higher rates over a given period (Chapman et al., 2011; Rumberger, 2011). While this article replicates previously documented racial graduation gaps, an exploration of more granular categories (in this case the interaction of gender, race, and income) reveals an unusual case of White students falling behind in attainment.[1]

We draw upon a large literature on the predictors of graduation or dropout to inform the choice of mediating factors. The contributions of family income, school, and neighborhood characteristics feature prominently in the literature from sociology to economics to developmental psychology (Autor et al., 2016; Coleman et al., 1966; Cook & Evans, 2000; Jencks, 1972; Rumberger, 2011).[2] We then consider cognitive ability because it is a well-established as a determinant of educational attainment (Heckman & Krueger, 2005; Paglin & Rufolo, 1990; Weinberger, 1999). The majority of literature finds a significant role of academic achievement on attainment outcomes (Rumberger, 2011). Finally, we examine socioemotional and behavioral dimensions at the individual level, given previous studies' findings on the importance for attainment of engagement and attitudes as measured by misbehavior, self-esteem, participation, expectations, and goals (Finn, 1989; Fredricks et al., 2004). While these aspects have long been studied in fields such as sociology and developmental psychology, our examination of socioemotional and behavioral dimensions reflect the shift in economics from a unidimensional focus on cognitive abilities toward the contributions of a wider range of noncognitive factors (Bowles et al., 2001; Cunha et al., 2010; West et al., 2016).

The remainder of the article proceeds as follows. The next section outlines the sample construction using two national longitudinal surveys as well as the North Carolina administrative data. The third section summarizes racial gaps in graduation rates by student subgroup and focuses on the unexpected deficit low-income White females have relative to Black females. The fourth section investigates the contributions of household income. The fifth section examines the role of neighborhood and school contexts, while the sixth and penultimate section provides evidence on the contribution of various cognitive and socioemotional or behavioral measures such as substance use. The final section concludes with a discussion of our findings.

Data

Our analyses use multiple longitudinal and administrative data sources. We follow Heckman and LaFontaine (2010) and Murnane (2013) in using two recent longitudinal data sets to capture U.S. graduation patterns: NLSY97 and Add Health. An advantage of these longitudinal surveys in addition to geographic coverage is the availability of measures on family context, behaviors, and attitudes that are not typically present in administrative data.[3] In contrast, administrative statewide records allow us to study a large population of students and characterize in detail the schools and peers these students were exposed to throughout primary and secondary schooling. We therefore supplement national-level surveys with statewide data from the NCERDC.

NLSY97

The NLSY97 sample comprises 8,984 individuals who were between the ages of 12 and 17 years during the initial interview in 1997. The sample follows these individuals through the 2015 survey wave, when respondents were 30 to 36 years old. We construct student groups at the intersection of race, gender, and economic disadvantage as measured by free or reduced-price lunch (FRL) eligibility, which occurs

when an individual's household income is at or below 185% of the federal poverty guideline. We use ethnicity data to define non-Hispanic White and African American students and code high school graduation based on students' reporting of enrollment status and the highest degree received during each interview. The sample excludes students who graduated with a GED, in keeping with official measures of high school graduation and previous studies documenting racial gaps in attainment.[4]

In addition to high school graduation, we rely on an array of survey questions to document students' academic achievement, behaviors, and attitudes. Cognitive scores derive from the Armed Services Vocational Aptitude Battery (ASVAB), a multiple-choice test covering aptitudes ranging from arithmetic reasoning to paragraph comprehension. Other school-based measures include tardies, absences, and grade retention. Family characteristics include a respondent's household income, while behavioral attributes cover tobacco, alcohol, marijuana, and other illicit drug use. To create the final sample, we first limit to the 8,768 students that complete eighth grade with nonmissing high school graduation outcomes.[5] Of these respondents, approximately one quarter were neither White nor African American. We then restrict to White or Black females with nonmissing family income data, before limiting the analytic sample to individuals who are low-income.

Add Health

We supplement the NLSY97 sample with Add Health, which surveyed students of a similar age and at a similar time. The initial wave of Add Health in 1994–1995 involved in-home interviews with 20,745 adolescents in Grades 7 to 12, when they were largely between the ages of 12 and 18. Add Health then followed these respondents through the latest Wave IV interview round in 2007–2008, when the remaining 15,701 participants were between 24 to 32 years old. The similarities between these two national data sets also extend to the timing of follow-up interviews. As such, this sample has a similar age distribution in any given year as the NLSY97.

We construct student groups using sociodemographic information taken in Wave I. A combination of parent-reported household income and household size determine the ratio of income to the federal poverty threshold. As with the NLSY97, we designate respondents as low-income if they are eligible for FRL. Attainment data comes from a Wave IV survey question on high school graduation status. Consistent with the NLSY97, we define graduation as finishing high school with a diploma instead of receiving a GED and exclude respondents who ever received a GED from the sample.

Add Health records a number of cognitive, socioemotional, and behavioral covariates. Achievement scores come from the Add Health Picture Vocabulary Test, an abridged version of the Peabody Picture Vocabulary Test with age-standardized scores. School variables include an indicator for ever repeating a grade, and the number of excused and unexcused absences. We expand the set of socioemotional measures to include self-esteem and expectations taken during the initial interview. The former derives from a modified Add Health version of the Rosenberg Self-Esteem Inventory.[6] The latter comes from survey questions eliciting participants' interest in and probabilities of going to college.[7] The final set of covariates capture individuals' tobacco, alcohol, and illicit substance use during Wave I. The sample restriction process is analogous to that of the NLSY97. We restrict to students who were in grade 9 or earlier during Wave I to ensure that baseline covariates are measured in earlier schooling years. These 7,069 observations decrease by nearly one fifth after dropping respondents who were not White or African American. We further exclude respondents who were male or missing income data. The final analytic sample restricts to only low-income Black or White females.

NCERDC

We complement these two national longitudinal surveys with administrative data on the universe of North Carolina high school students. The NCERDC allows us to follow ninth-grade students in 2008 to 2011 through high school graduation. We construct student subgroups using race, gender, and FRL eligibility from student roster files.[8] Notably, FRL eligibility can fluctuate across time as parental incomes change or because of administrative errors. To ensure we are distinguishing families that are persistently economically advantaged or disadvantaged from families experiencing income shocks, we rely on the longitudinal nature of the data and compile FRL eligibility for students going as far back as third grade. Our measure for low-income is students who were *ever* eligible for FRL, whereas higher income refers to students who were *never* eligible.

The outcome variable is whether a student graduated high school within 4 years of beginning ninth grade.[9] Specifically, a ninth grader in 2008 is matched to 2011 graduation files and 2008 to 2011 dropout and school exit data to ensure a comprehensive view of their manner of exit.[10] School-level attributes include whether the high school is classified as urban, suburban, or rural.[11] School rosters permit us to construct grade-level student composition by gender, race, and economic disadvantage. The final analytic sample spans the 2008–2011 ninth-grade cohorts. We exclude students who exit for reasons that are exempt from statewide graduation rate calculations, as well as students with missing exit reason or timing.[12] We retain unique individual observations by only including the first time an individual was observed in a ninth-grade cohort. The full sample comprises 406,349 students, while the restricted sample with only White and African American students with nonmissing sociodemographic data includes 342,246 individuals.

TABLE 1
Graduation Rates

Group	Add Health	NLSY97	North Carolina
Overall	78.0%	77.7%	81.1%
Overall racial gap			
White	81.5%	80.2%	84.9%
Black	72.8%	69.9%	76.4%
Black-White gap	−8.7	−10.3	−8.5
Racial gap among males			
White	79.6%	78.3%	82.5%
Black	65.0%	64.0%	70.5%
Black-White gap	−14.6	−14.3	−12.0
Racial gap among females			
White	83.6%	82.2%	87.4%
Black	81.5%	76.1%	82.4%
Black-White gap	−2.1	−6.2	−5.1
Racial gap among low-income males			
White	66.1%	61.2%	67.4%
Black	56.0%	51.3%	67.3%
Black-White gap	−10.1	−9.9	−0.1
Racial gap among low-income females			
White	72.3%	66.1%	74.9%
Black	78.3%	72.1%	80.3%
Black-White gap	6.0	6.0	5.4

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997; GED = general educational development. The Add Health sample of 7,069 respondents includes those first observed no later than ninth grade and have non-missing graduation data in Wave IV. Racial gap estimates furthermore exclude non-White or non–African American respondents and observations with missing family income or household size data. Results are weighted using GSWGT4_2 to account for nonrandom attrition from sample. Completion outcomes observed in Wave IV excludes GED recipients. The NLSY97 sample of 8,768 respondents includes those that completed eighth grade, with nonmissing highest grade completed and high school graduation data by the 2015 survey wave. Results are weighted using the Round 1 sampling panel weight (R1236201). Completion outcomes observed in the 2015 survey wave excludes GED recipients. The North Carolina sample includes students who enroll in the ninth grade for the first time in 2008 to 2011. The outcome is a dichotomous variable for whether a student graduates high school within 4 years. The full North Carolina sample includes 406,349 observations. Note that we do not weight the Add Health or NLSY97 estimates in the fourth panel, since we condition on household income. See text for discussion.

An Exception to the Black-White Attainment Gap

The U.S. high school graduation rate is a barometer for both the performance of schools as well as the population's skill level (Heckman & LaFontaine, 2010). As such, the overall rate as well as the graduation outcomes of student subgroups by race and gender is the subject of ample scholarly attention. We begin by corroborating the overall graduation rates reported in this literature using our three analytic samples (Table 1). The first panel shows broad consistency across the two national samples with rates that closely adhere to previous estimates. The Add Health and NLSY97 samples report that 78% and 77.7% of students earn a high school degree, respectively.[13] Likewise, 81.1% of students in North Carolina public schools graduate within 4 years of beginning ninth grade, which is within 0.3 percentage points of official 4-year graduation rate reports.[14] Any disparities between national- and state-level estimates can depend on differences in student composition, cohorts, and choice of attainment measure.[15]

We proceed next to examine Black-White disparities.[16] The graduation gap ranges from 8.5 to 10.3 percentage points across the three samples, consistent with gaps reported in Heckman and LaFontaine (2010). Research on women's attainment advantage over men and the increasing divergence between the academic trajectories of Black boys and girls motivates further segmenting the sample by gender (McDaniel et al., 2011; Saunders et al., 2004). The racial gap among boys of 12.0 to 14.6 percentage points is consistently larger than the female racial gap. Even then, we observe the empirical regularity of White students graduating high school at higher rates than their Black peers across samples and subgroups.

Previous studies examined how graduation rates vary by the interaction of gender and race, but less attention is given to how attainment outcomes vary by socioeconomic background (Murnane, 2013).[17] Research on the Black-White achievement gap that controls for socioeconomic status find that the latter accounts for anywhere between a quarter up to 85% of disparities at various points of the educational experience (Clotfelter et al., 2009; Fryer & Levitt, 2004, 2006; Murnane et al., 2006; Reardon & Robinson, 2008).[18] Moreover, there exists substantial heterogeneity by gender conditional on income (Chetty et al., 2018). A growing body of evidence finds that the effects of socioeconomic status on development and academic performance vary significantly across gender groups (Autor et al., 2016, 2019; Bertrand & Pan, 2013; Owens, 2016). This motivates a closer exploration of racial gaps in graduation for specific gender and income subgroups.

We use income as a proxy measure for socioeconomic status and report racial differences in attainment among the economically disadvantaged in the final panel of Table 1.[19] Among boys, Black students lag significantly behind girls in graduation across both national samples and are closer to parity in the North Carolina data set. Economically disadvantaged girls, on the other hand, reverse the traditional narrative of Black students falling behind in high school graduation. Low-income White girls graduate at a rate between 5.4 and 6 percentage points *lower* than their Black peers.[20] This unexpected graduation deficit for White girls is not limited to particular samples and birth cohorts. Our national and statewide administrative samples match the characteristics of similar samples used in the literature and the official statistics published by the state board of education. Furthermore, each of

TABLE 2
Family Incomes Among Low-Income Females

Variable	Add Health		NLSY97	
	White	Black	White	Black
Graduation rate	72.3%	78.3%**	66.1%	72.1%*
Household income				
Multiples of poverty level, %	113%	89%***	108%	82%***

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997. Add Health sample includes all White or African American low-income females who were in ninth grade or earlier during Wave I. High school graduation is defined using attainment data in Wave IV. NLSY97 sample includes all White and non-Hispanic Black respondents who completed eighth grade. Household incomes are computed using household income, size, and federal poverty thresholds in 1994.
*$p < .1$. **$p < .05$. ***$p < .01$.

these data sets find graduation gaps that are consistent in sign and magnitude across birth cohorts spanning the late 1970s through the mid-1990s.

The question arises, then, of why low-income White females lag behind in high school graduation relative to Black females. We begin our inquiries by examining three sets of possible explanations for our findings. First, we verify that our results are not explained by higher household income among Black females and that results are robust to alternative definitions of economic disadvantage. Second, we explore the role played by a student's family, neighborhood, and school context, such as the urbanicity of her neighborhood or the peer composition of her high school. Third, we examine whether low-income Black women graduate at higher rates because of greater endowments of cognitive and socioemotional skills as measured by school retention and delinquency, risky behaviors, and related attributes that predispose them to higher rates of high school graduation. After scrutinizing these factors, we consider reasons for why attainment patterns among low-income females differ markedly from well-established patterns of attainment, including among similarly disadvantaged males.

The Role of Economic Disadvantage

We first consider the possibility that among students who are economically disadvantaged, Black students may still have higher average family incomes when using an alternative, more fine-grained measure, and this can contribute to their better attainment outcomes. To the contrary, we show that low-income Black females come from households earning significantly *less* on average than White females. Both national longitudinal surveys find that low-income White females come from households at 108% to 113% of the federal poverty limit, while Black females come from households earning just 82% to 89% (Table 2).[21] Black students succeed in graduating high school in spite of their income levels, not because of them.

A related concern is that our binary income classification masks important heterogeneity within the low-income female population; namely, the magnitude of the racial attainment gap for individuals from the most impoverished families relative to those near and above the 185% threshold is unclear. We therefore examine the sensitivity of our results across the full spectrum of socioeconomic backgrounds. Figure 1 plots a LOWESS-smoothed graph of the high school graduation rate for both the NLSY97 and Add Health across the ratio of income to poverty level, with a vertical marker for the 185% cutoff. The figure shows that the graduation advantage among economically disadvantaged Black females is robust to the choice of income cutoff. Black females have a higher graduation rate than White females over a broad income range, from zero to over two times the federal poverty level. Across both samples, White and Black females in households earning between four and five times the poverty level have similar graduation rates. These results underscore the importance of parsing racial gaps among females into its parts: differences in the distribution of household incomes for White and Black females, and differences in graduation rates at a given household income level.

Our analysis of North Carolina administrative data further supports the view that low-income White students underperform despite coming from families that are more economically privileged than their Black peers. The present categorization of economic disadvantage includes students who were ever eligible for FRL between 8th and 10th grades. This groups students experiencing transitory negative income shocks around the 185% threshold with those experiencing more permanent shocks and students misclassified as low income. Appendix Table A2 divides this lower income population into those who were consistently FRL eligible during elementary and middle school (bottom income level) versus those who had mixed eligibility records over time (middle income level). We supplement these results with median incomes of the student's residential block group from the American Community Survey. We find that Black women consistently graduate at higher rates than White women
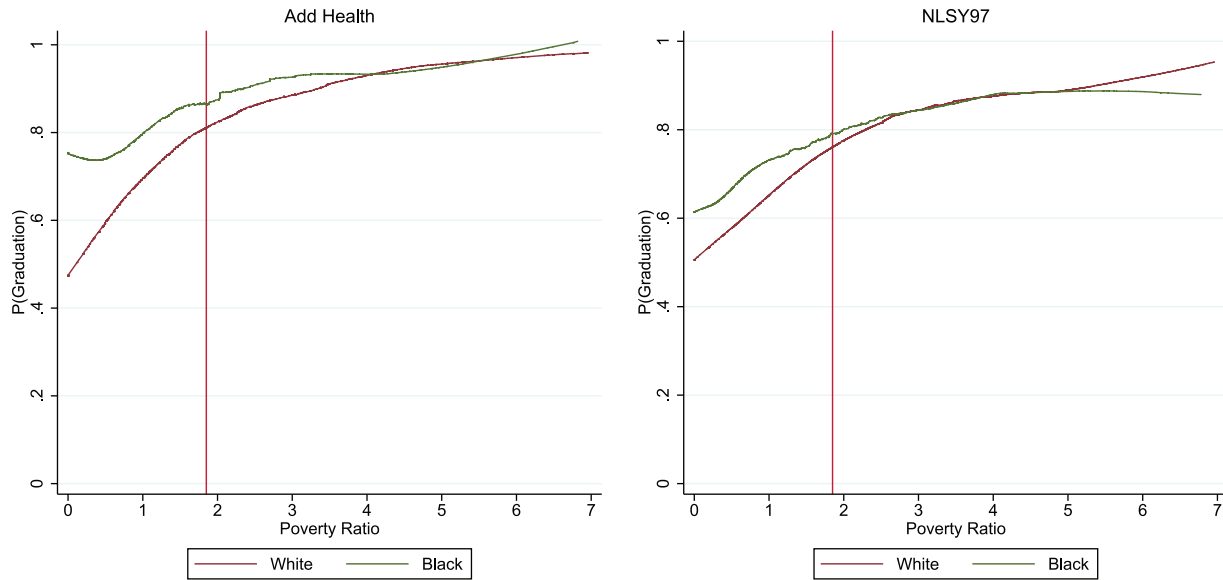
FIGURE 1. *Graduation rate by ratio of income to poverty.*
Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997. Locally weighted smoothed relationship between family poverty ratio and probability of graduating high school. The Add Health and NLSY97 samples include all White or African American respondents with nonmissing educational attainment data.

TABLE 3
North Carolina—Neighborhood Factors

Variable	(1)	(2)	(3)
Black female	0.054*** (0.006)	0.097*** (0.005)	0.099*** (0.008)
Black female × mixed block group[a]			−0.004 (0.010)
Black female × rural block group[b]			−0.001 (0.013)
Additional controls			
Block group fixed effects	N	Y	Y
Observations	76,599	76,599	76,599

Note. The full sample includes economically disadvantaged ninth-grade Black and White female students during 2008–2011. Standard errors clustered at the school level.
[a]Housing units in block group are classified as a mix of urban and rural in the 2010 Census. [b]All housing units in block group are classified as rural in the 2010 Census.
*p < .10. **p < .05. ***p < .01.

across all income categories. The lower graduation rates of economically disadvantaged White females are most pronounced among the worst-off, even though White students reside in block groups with family incomes that average $6,400 higher. These results from North Carolina echo findings in Figure 1 that the lagging graduation rates of White students is starkest among the most economically disadvantaged families.[22]

The Role of Neighborhood and School Contexts

We rely on geocoded data from North Carolina to explore whether the sorting of students by race into urban or rural neighborhoods contributes to the graduation patterns we observe. Population density can be associated with varying levels of social service provision, labor market opportunities, and the quality of education and related support services that may influence graduation outcomes. Table 3 shows estimates from linear probability models accounting for students' residential block groups. The baseline Black attainment edge is 5.4 percentage points in North Carolina, with low-income Black females graduating at 80.3%, while their White counterparts lag at 74.9% (Table 1). This disparity increases with the addition of block group fixed effects, suggesting that the phenomenon we observe is pervasive even within fine-grained neighborhoods.[23] Notably, we find no evidence that these patterns vary by the rurality of students' residential locations.

We next examine students' school and classroom environments in contributing to observed graduation disparities.

TABLE 4
North Carolina—School Sorting and Student Composition

Variable	(1)	(2)	(3)	(4)
Black female	0.054*** (0.006)	0.081*** (0.005)	0.083*** (0.005)	0.107** (0.043)
Black female × share of disadvantaged White students				0.103* (0.060)
Black female × share of disadvantaged Black students				0.007 (0.059)
Black female × share of advantaged White students				−0.131** (0.053)
Black female × share of advantaged Black students				−0.092 (0.146)
School fixed effects	N	Y	N	N
School × year fixed effects	N	N	Y	Y
Observations	76,599	76,599	76,599	76,599

Note. The full sample includes economically disadvantaged ninth-grade Black and White female students during 2008 to 2011. The omitted group is the share of non-White and Black students in the individual's ninth-grade class. Standard errors are clustered at the school level.
*$p < .10$. **$p < .05$. ***$p < .01$.

One potential explanation for the gap is that low-income Black females systematically attend schools that graduate students at higher rates. These schools could have programs that better serve students at risk of dropping out or credit recovery programs that affect the probability of completion. We assess whether increasing racial segregation and sorting by school attributes play a role by augmenting the base model with school fixed effects. The first column in Table 4 shows the same unadjusted graduation advantage of 5.4 percentage points among poorer Black females. Similar to the neighborhood analysis, the graduation gap increases in the school fixed effects model to 8.1 percentage points. This suggests that the graduation edge of low-income Black females persists within schools. Further restraining the racial comparison to within school and year combinations shows that low-income Black females graduate at 8.3 percentage points higher than their White peers.

Finally, we interact the coefficient of interest with school-level student composition variables to assess the extent to which the graduation gap's magnitude depends on the prevalence of each sociodemographic subgroup. Column (4) shows that the Black female attainment advantage is increasing in the share of economically disadvantaged White students and decreasing in the share of better-off White students. The gap is not significantly affected by other student groups. The association between graduation rates and peer composition within the school setting suggests a social dimension to these patterns. Social networks and its influence on risky behaviors and attainment may merit closer attention (Kawaguchi, 2004; Lundborg, 2006).

The Role of Cognitive and Socioemotional Skills

We shift from neighborhood and school-level attributes to individual characteristics. We start by examining the role that cognitive factors play in explaining the attainment gap and follow with a detailed analysis of how socioemotional and behavioral attributes account for observed disparities.

Cognitive Skills

It is well documented that cognitive skills shape educational attainment, labor market outcomes, and other dimensions of individual well-being (Hanushek & Woessmann, 2008; Heckman, 2008). The first step in our analysis explores the variation in this determinant of schooling to see whether higher graduation rates among low-income Black females is explained by superior cognitive ability. To do so we focus on standardized test scores and average student GPAs on a 4-point scale.[24] Even though many social science researchers use achievement test scores, GPA, and IQ to capture dimensions of cognitive functioning, it is increasingly recognized that personality traits and other noncognitive skills play a role in the formation of grades (see, e.g., Heckman & Kautz, 2012). While we use GPA as a cognitive measure, our qualitative results are robust to excluding GPA as a measure of cognitive ability.

Table 5 summarizes test scores and grades. Low-income Black females have test scores that are up to two thirds of a standard deviation lower, and eighth grade GPAs that are between 0.1 and 0.14 points lower than economically disadvantaged White females.[25] These means suggest that conditioning on test scores and GPA would only reinforce the Black female graduation advantage, under the assumption that they are positively correlated with graduation outcomes. Table 6 regresses high school graduation on being a low-income Black female with baseline controls, with additional models augmented with household income and cognitive attributes. Column (1) shows that poorer Black females in Add Health and NLSY97 graduate, respectively, at 6.6 and 6.4 percentage points higher than their White counterparts. When we include household income, the attainment edge of economically disadvantaged Black women *increases*. The same occurs when conditioning on test scores and GPA, suggesting that equalizing these measures of cognitive ability further increases the graduation advantage of low-income Black females who lag in these positive correlates

TABLE 5
Cognitive Skills Among Low-Income Females

Variable	Add Health		NLSY97	
	White	Black	White	Black
Graduation rate	72.3%	78.3%**	66.1%	72.1%*
Cognitive measures				
Test scores	0.03	−0.64***	0.27	−0.30***
GPA	2.72	2.62*	2.82	2.68***

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997; GPA = grade point average. Add Health sample includes all White or African American low-income females who were in ninth grade or earlier during Wave I. High school graduation is defined using attainment data in Wave IV. NLSY97 sample includes all White and non-Hispanic Black respondents who completed eighth grade. Test scores from the Picture Vocabulary Test in Add Health and the Armed Forces Qualifications Test scores in the NLSY97 are standardized to *z* scores. GPA from Add Health is the average math and English GPA on a 4-point scale.
*$p < .1$. **$p < .05$. ***$p < .01$.

of completion. The findings are robust to the use of multiple imputation. Appendix Table A4 shows a graduation advantage among low-income Black females that are similar in magnitude after accounting for missing values.

Socioemotional and Behavioral Attributes

A substantial body of research relates the role of non-cognitive factors such as behaviors and attitudes to educational and labor market outcomes (Bowles & Gintis, 1976; Farkas, 2003; Heckman & Rubinstein, 2001; Rumberger, 2011). The capacity of these characteristics to influence earnings independent of traditional cognitive measures is well-recognized (Almlund et al., 2011; Borghans et al., 2014; Jencks, 1979; Murnane et al., 2001). Research increasingly documents the relevance of such inputs for drop out and graduation (Alexander et al., 1997; Aucejo, 2015; Heckman et al., 2006).[26]

We examine the role of socioemotional and behavioral attributes for the observed graduation gap. Survey items taken during initial waves, when students were at the beginning of their high school education or in middle school, cover a range of schooling and health behaviors and socioemotional attributes. We follow previous literature in the selection of these variables to capture ability outside of cognitive domains. Measures from school records include retention, absences, tardies, or general delinquency (Jackson, 2018). We rely on risky behaviors during childhood such as the use of alcohol, tobacco, marijuana, and illicit drugs that have been shown to predict academic outcomes (Crosnoe, 2006; Cunha & Heckman, 2008). We use the Rosenberg Self-Esteem Scale in Add Health due to its correlation with the traditional Big Five typology used as a long-standing socioemotional measure (Almlund et al., 2011). Remaining measures including self-reported desire and likelihood of attending college and life experiences such as bullying approximate individual expectations and hardships.

Summary statistics in Table 7 show that White females lag behind across a number of these measures. While low-income Black females are more likely to be retained at least one grade before entering high school, they have fewer absences and score one sixth lower on a delinquency index than low-income White females. An even starker contrast is apparent when examining measures of substance use. Low-income Black females are significantly less likely to have reported ever using tobacco, alcohol, marijuana, and other illicit drugs during adolescence. In addition to striking differences at the extensive margin, results not shown here document that economically disadvantaged Black females who have been initiated into these substances also use them at lower frequency. These racial gaps among the focal group of disadvantaged females echo findings from a broad literature spanning epidemiology and public health that Black adolescents record lower alcohol, tobacco, and illicit drug consumption than their non-Hispanic White peers (Keyes et al., 2015; Swendsen et al., 2012).[27] Black females furthermore profess greater intention as well as likelihood of attending college compared to White peers.

Table 8 examines the extent to which these measures account for the observed Black female attainment advantage by reporting coefficients from linear probability models regressing high school graduation on being a low-income Black female, baseline controls, and socioemotional or behavioral attributes. The first specification confirms low-income Black females in both longitudinal surveys graduate at rates that are more than 6 percentage points higher than White females.

Specifications in columns 2 to 7 consider the role of socioemotional and behavioral measures. Accounting for absenteeism and retention, substance use, self-esteem, and college expectations in the Add Health sample leaves a graduation gap of 3.2 percentage points that is no longer significantly different across racial groups. Similarly, adding controls for absenteeism, delinquency, retention, substance use, and hardship in the NLSY97 sample decreases the graduation gap to a nonsignificant 2.7 percentage points.[28] A consistent feature of these models is an attenuated racial graduation disparity when accounting for the use of tobacco, alcohol, marijuana, cocaine, or other illicit drugs during early adolescence.

We estimate the contribution of individual socioemotional and behavioral measures by relying on the Gelbach decomposition technique. This approach, based on the sample omitted variables bias formula, addresses the issue of sequencing sensitivity to render estimates invariant to the order in which variables are added to the specification (Gelbach, 2016). The relevance of a covariate derives from two factors when we apply this technique to the Black-White graduation gap: (1) Black-White differences in this covariate after partialing out all other explanatory variables in the base

TABLE 6
Income, Cognitive Measures, and High School Graduation

Variable	1	2	3
Add Health (*n* = 886)			
Black female	0.066* (0.034)	0.098*** (0.035)	0.135*** (0.033)
Additional controls			
Household poverty ratio	N	Y	N
Test scores and GPA	N	N	Y
NLSY97 (*n* = 872)			
Black female	0.064* (0.037)	0.104*** (0.037)	0.165*** (0.032)
Additional controls			
Household poverty ratio	N	Y	N
Test scores and GPA	N	N	Y

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY = National Longitudinal Survey of Youth 1997; GPA = grade point average; SES = socioeconomic status. The table shows coefficients corresponding to linear probability models regressing high school graduation outcomes on an indicator for being a low-income Black female, birth year, and additional covariates. The Add Health sample includes low-SES White or African American respondents who were in ninth grade or earlier during Wave I. Standard errors are clustered at the school level. The NLSY97 sample includes low-SES White and non-Hispanic Black respondents who completed eighth grade. Standard errors are clustered at the NLSY97 primary sampling unit and stratum. The coefficient for the Black female indicator varies from the summary statistics because these specifications adjust for birth year and drop respondents with missing values for the additional controls.
*$p < .1$. **$p < .05$. ***$p < .01$.

regression and (2) the extent to which the covariate is correlated with educational attainment.[29] Results from the decomposition are consistent estimates of the conditional contribution of each variable.

We find that behavioral and personality proxies for students' noncognitive skills collectively explain between 51% and 59% of the Add Health and NLSY97 samples, respectively (Table 9). Factors such as absenteeism, delinquency, retention, and self-esteem play no significant role.[30] The largest contributor accounting for between 34% and 38% of the graduation gap between low-income Black and White females in both samples is adolescent substance use. Economically disadvantaged White females are significantly more likely to use every drug: 20 to 23 percentage points more likely to use tobacco, 12 to 17 percentage points more likely to use alcohol, and 6 to 7 percentage points more likely to use marijuana. Even more alarming is the finding that low-income White women are overwhelmingly more likely to use illicit drugs like cocaine, heroin, and methamphetamines than Black women. No more than 2% of low-income Black females use these illicit drugs in our samples, while between 9% and 11% of low-income White females do so across our two samples.

Discussion

Our findings show that the early onset of substance use among low-income White females predicts lower attainment. We cannot rule out the possibility of omitted variables, as hard-to-observe factors that induce greater substance use may also adversely affect schooling attainment. Substance use can exert an influence on educational attainment on its

own, or its explanatory power can derive from correlated factors. It has been documented, for instance, that mental health issues and substance use often co-occur among adolescents, which motivates studies of their joint consequences for educational attainment (Breslau et al., 2011; Macleod et al., 2004). A growing body of research also finds that mental health problems are directly associated with lower high school completion (Breslau et al., 2008; Evensen et al., 2016; Kessler & Saunders, 1995). In light of previous evidence, we verify that our findings are robust to the inclusion of mental health variables in both national longitudinal surveys. We show that these factors do not have a sizable impact on the Black-White attainment gap when included on their own or concurrently with substance use covariates (Appendix Table A6). Note that this does not preclude an association between mental health factors and educational attainment and only suggests that these variables play a lesser role than substance use for racial disparities in attainment.

Our results call to attention literature showing the effect of substance use on schooling outcomes. Chatterji (2006) documents a relationship between illicit drug use and attainment while Marie and Zölitz (2017) show that substance use lowers academic performance.[31] Beyond academic outcomes, tobacco, alcohol, and drug use also have negative ramifications for health, spanning from cardiovascular disease to increased injuries and mortality (Carpenter & Dobkin, 2009; Hall & Degenhardt, 2009). Recent findings of rising mortality and morbidity among less-educated non-Hispanic White adults emphasize the role of drug overdoses, suicides, and alcohol-related liver mortality in disproportionately affecting this sociodemographic group, with evidence suggesting worsening labor market opportunities with accompanying

TABLE 7

Socioemotional Attributes Among Low-Income Females

Variable	Add Health		NLSY97	
	White	Black	White	Black
Graduation rate	72.3%	78.3%**	66.1%	72.1%*
School records				
Ever retained	25%	38%***	17%	26%***
Days absent			6.35	5.20**
Excused absences: Never	9%	15%***		
Excused absences: 1 to 2 times	25%	37%***		
Excused absences: 3 to 10 times	46%	39%**		
Excused absences: 10+ times	20%	8%***		
Unexcused absences	0.59	0.38***		
Days tardy			2.34	2.51
Delinquency index			1.20	0.97**
Substance use (% ever)				
Tobacco	66	46***	52	29***
Alcohol	55	38%***	46	34***
Marijuana	22	16**	23	16***
Illicit drugs	11	1***	9	2***
Cocaine	4	0***		
Other	11	0***		
Self-esteem	12.44	10.49***		
College				
Want to go	4.36	4.56***		
Likelihood of going	3.98	4.19***		
Hardship				
Experienced hard times			12%	8%
Ran away from home			15%	11%
Bullied			22%	23%

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997. The Add Health sample includes all White or African American respondents who were in ninth grade or earlier during Wave I. NLSY97 sample includes all White and non-Hispanic Black respondents who completed eighth grade.
*$p < .1$. **$p < .05$. ***$p < .01$.

cumulative disadvantage for this group (Case & Deaton, 2017). While our study timeline stops far short of later life morbidity and mortality, it prompts questions about the prelabor market origins of such patterns among the low-income White population.

The question remains of why the Black-White graduation gap is only reversed for low-income females instead of males of similar economic standing. It may be that economically disadvantaged White women fare particularly poorly in educational attainment relative to White men, or Black women perform better than predicted. The former possibility has an analogous result in the recent "deaths of despair" literature, which documents a higher mortality increase for non-Hispanic White women than men (Gelman & Auerbach, 2016; Woolf et al., 2018). Causes for rising female death rates include drug overdoses linked to the epidemic of prescription opioids, smoking- and obesity-related illnesses, and suicide (Astone et al., 2015; Woolf et al., 2018).[32] The

observation that these patterns are particularly salient among middle-aged White women in the southern United States suggests closer study is warranted for understanding afflictions that are specific to this population (Gelman & Auerbach, 2016).

At the same time, the Black-White reversal in graduation rates may be attributed to relatively better performance among low-income Black girls. A growing literature documents academic resilience among economically disadvantaged minority populations (Catterall, 1998; Strand, 2014). Researchers are also repeatedly documenting a gender difference in sensitivity to family disadvantage, with boys' behavioral and academic outcomes more adversely affected than girls' (Autor et al., 2019; Bertrand & Pan, 2013).[33] Taken together, these strands of research suggest that Black females from low-income households may fare better than their brothers in ways that contribute to the graduation patterns we observe.

TABLE 8
Socioemotional Measures and High School Graduation

Variable	1	2	3	4	5	6	7
Add Health (*n* = 886)							
Black female	0.066* (0.034)	0.056 (0.035)	0.030 (0.032)	0.045 (0.035)	0.047 (0.032)		0.032 (0.033)
Additional controls							
Absenteeism and retention		Y	N	N	N		Y
Substance use		N	Y	N	N		Y
Self-esteem		N	N	Y	N		Y
Expectations		N	N	N	Y		Y
NLSY97 (*n* = 872)							
Black female	0.064* (0.037)	0.047 (0.034)	0.017 (0.036)			0.047 (0.034)	0.027 (0.034)
Additional controls							
Absenteeism, delinquency, and retention		Y	N			N	Y
Substance use		N	Y			N	Y
Hardship		N	N			Y	Y

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY = National Longitudinal Survey of Youth 1997; SES = socioeconomic status. The table shows coefficients corresponding to linear probability models regressing high school graduation outcomes on an indicator for being a low-income Black female, birth year, and additional covariates. The Add Health sample includes low-SES White or African American respondents who were in ninth grade or earlier during Wave I. Standard errors are clustered at the school level. There are 886 observations across all specifications. The NLSY97 sample includes low-SES White and non-Hispanic Black respondents who completed eighth grade. Standard errors are clustered at the NLSY97 primary sampling unit and stratum. There are 872 observations across all specifications. The coefficient for the Black female indicator varies from the summary statistics because these specifications adjust for birth year and drop respondents with missing values for the additional controls.
*p < .1. **p < .05. ***p < .01.

TABLE 9
Decomposition

	Add Health		NLSY97	
Variable	Contribution	Share of Gap (%)	Contribution	Share of Gap (%)
Absenteeism and retention	0.002	3.2	−0.005	−7.1
Substance use	−0.022**	−33.7	−0.025**	−38.2
Self-esteem	−0.002	−3.1		
Expectations	−0.012**	−17.8		
Hardship			−0.009	−13.2
Total	−0.034*	−51.4	−0.038**	−58.5

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997. The Add Health sample includes all White or African American respondents who were in ninth grade or earlier during Wave I. NLSY97 sample includes all White and non-Hispanic Black respondents who completed eighth grade. The attainment advantage in the Add Health base model of 6.6 percentage points decreases to 3.2 percentage points in the full sample.
*p < .1. **p < .05. ***p < .01.

Conclusions

Black-White disparities in educational achievement and attainment are a well-documented empirical regularity (Fryer & Levitt, 2006; Murnane et al., 2006; Reardon & Robinson, 2008). However, disaggregating by income and gender reveals a striking finding: high school graduation rates among low-income White females are consistently 5 to 6 percentage points lower than their Black female peers. Data from two national longitudinal surveys and administrative records from the state of North Carolina verify that these patterns persist despite economically disadvantaged Black females averaging lower household incomes and standardized test scores. As such, the achievement gap we observe between Black and White students does not translate to an analogous disparity in high school graduation rates.

To account for White students' relative underperformance in high school graduation, we examine disparities in cognitive skills and a range of socioemotional and behavioral

attributes spanning school delinquency, substance use, and self-esteem. Accounting for standardized test scores and grades magnifies the attainment advantage of poorer Black females. In contrast, behavioral factors markedly close the gap, with substance use explaining approximately one third of the overall disparity across both national samples. Low-income Black females average significantly lower average rates of tobacco, alcohol, marijuana, and illicit drug use than their White peers.

The evidence shows that socioemotional proxies such as the early onset of risky health behaviors are relevant factors for high school graduation. Closer scrutiny of these inputs is warranted for improving graduation rates across student groups, particularly given the rising labor market returns to schooling. Two thirds of earnings inequality in the United States from 1980 to 2005 is attributable to greater schooling premiums (Autor, 2014; Goldin & Katz, 2007), such that the difference between dropping out and completing high school with the option to pursue higher education is increasingly stark.[34]

A key implication of this study is that examining attainment inequality from an exclusively racial or gender lens obscures the trends that are distinct to groups at the intersection of sociodemographic categories. Similar levels of economic disadvantage can materialize in different ways for each group to put downward pressure on high school graduation rates. By examining patterns at the intersection of race, gender, and income, we are able to show that disadvantaged Black females challenge the conventional narrative on racial graduation gaps. They maintain lower rates of substance use and school-related delinquencies despite residing in some of the most underprivileged neighborhoods and scoring below White peers on standardized tests. Their unexpected edge in educational attainment suggests that efforts to improve graduation outcomes can focus on Black females' resilience to disadvantage and challenges disproportionately faced by their White peers. More broadly, a closer examination of students in granular sociodemographic categories can inform tailored policies to improve high school graduation outcomes.

Appendix

TABLE A1

Graduation Rates—Imputed Household Income

Group	Add Health (%)	NLSY97 (%)	North Carolina (%)
Low-income females			
White	72.1	64.8	74.9
Black	77.6	71.1	80.3
Black-White gap	5.5	6.3	5.4

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997; SES = socioeconomic status. The Add Health sample of 1,052 respondents includes low-SES White and Black female respondents first observed no later than ninth grade with nonmissing graduation data in Wave IV. The NLSY97 sample of 1,095 White and Black female respondents includes those that completed eighth grade, with nonmissing highest grade completed and high school graduation data by the 2015 survey wave. The North Carolina sample remains the same as Table 1.

TABLE A2

North Carolina—Alternative Definitions of Disadvantage

	FRL[a]		
Group	Share (%)	Graduation Rate (%)	Block Group Income($)[b]
Middle income level			
Black female	5	80	41,527
White female	6	77	46,734
Bottom income level			
Black female	9	80	38,059
White female	5	72	44,458

Note. FRL = free or reduced-price lunch. Sample includes all students observed in ninth grade between 2008 and 2011.
[a]Middle income level is defined as students who were both eligible and ineligible for free and reduced lunch during the period they were observed in the data set. Bottom income level is defined as students who were always eligible for FRL. [b]Block group median income come from the American Community Survey 2013 5-year sample.

TABLE A3

Add Health—Neighborhood and School Factors

Variable	1	2	3
Black female	0.066* (0.034)	0.078* (0.044)	0.127** (0.055)
Additional controls			
Neighborhood characteristics	N	Y	N
School fixed effects	N	N	Y
Observations	886	873	886

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; The Add Health sample includes all White or African American respondents who were in ninth grade or earlier during Wave I. Neighborhood block group characteristics including urbanicity, share of the population by racial group, average per capita income, and proportion of age 25+ year-old adults with at least a college degree.
*$p < .1$. **$p < .05$. ***$p < .01$.

TABLE A4

Income, Cognitive Measures, and High School Graduation

Variable	1	2	3
Add Health ($n = 1,052$)			
Black female	0.062* (0.033)	0.103*** (0.042)	0.131*** (0.033)
Additional controls			
Household poverty ratio	N	Y	N
Test scores and GPA	N	N	Y
NLSY97 ($n = 1,095$)			
Black female	0.059* (0.035)	0.159*** (0.043)	0.153*** (0.033)
Additional controls			
Household poverty ratio	N	Y	N
Test scores and GPA	N	N	Y

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997; GPA = grade point average; SES = socioeconomic status. These specifications use multiple imputation for missing observations. The table shows coefficients corresponding to linear probability models regressing high school graduation outcomes on an indicator for being a low-income Black female, birth year, and additional covariates. The Add Health sample includes low-SES White or African American respondents who were in ninth grade or earlier during Wave I. Standard errors are clustered at the school level. The NLSY97 sample includes low-SES White and non-Hispanic Black respondents who completed eighth grade. Standard errors are clustered at the NLSY97 primary sampling unit and stratum.
*$p < .1$. **$p < .05$. ***$p < .01$.

TABLE A5

Socioemotional Measures and High School Graduation

Variable	1	2	3	4	5	6	7
Add Health ($n = 1,052$)							
Black female	0.062* (0.033)	0.042 (0.033)	0.027 (0.032)	0.042 (0.033)	0.041 (0.032)		0.011 (0.032)
Additional controls							
Absenteeism and retention	N	Y	N	N	N		Y
Substance use	N	N	Y	N	N		Y
Self-esteem	N	N	N	Y	N		Y
Expectations	N	N	N	N	Y		Y
NLSY97 ($n = 1,095$)							
Black female	0.059* (0.035)	0.053* (0.032)	0.012 (0.034)			0.043 (0.032)	0.034 (0.031)

TABLE A5. (CONTINUED)

Variable	1	2	3	4	5	6	7
Additional controls							
Absenteeism, delinquency, and retention	N	Y	N			N	Y
Substance use	N	N	Y			N	Y
Hardship	N	N	N			Y	Y

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997; SES = socioeconomic status. These specifications use multiple imputation for missing observations. The table shows coefficients corresponding to linear probability models regressing high school graduation outcomes on an indicator for being a low-income Black female, birth year, and additional covariates. The Add Health sample includes low-SES White or African American respondents who were in ninth grade or earlier during Wave I. Standard errors are clustered at the school level. The NLSY97 sample includes low-SES White and non-Hispanic Black respondents who completed eighth grade. Standard errors are clustered at the NLSY97 primary sampling unit and stratum.
*p < .1. **p < .05. ***p < .01.

TABLE A6
Mental Health Factors and High School Graduation

Variable	1	2	3
Add Health (*n* = 886)			
Black female	0.066* (0.034)	0.060 (0.038)	0.039 (0.037)
Additional controls			
Mental health		Y	Y
Absenteeism and retention		N	Y
Substance use		N	Y
Self-esteem		N	Y
Expectations		N	Y
NLSY97 (*n* = 872)			
Black female	0.064* (0.037)	0.067* (0.035)	0.033 (0.033)
Additional controls			
Mental health		Y	Y
Absenteeism, delinquency, and retention		N	Y
Substance use		N	Y
Hardship		N	Y

Note. Add Health = National Longitudinal Study of Adolescent to Adult Health; NLSY97 = National Longitudinal Survey of Youth 1997. The table shows coefficients corresponding to linear probability models regressing high school graduation outcomes on an indicator for being a low-income Black female, birth year, and additional covariates. The Add Health sample includes all White or African American respondents who were in ninth grade or earlier during Wave I. Mental health factors include four Wave I variables on the frequency of crying, feeling moody, blue, or depressed. Standard errors are clustered at the school level. There are 886 observations across all specifications. The NLSY97 sample includes all White and non-Hispanic Black respondents who completed eighth grade. Mental health factors include three variables taken in 2000 on how often the respondent felt nervous, blue, or depressed over the previous month. Standard errors are clustered at the NLSY97 primary sampling unit and stratum. There are 872 observations across all specifications.
*p < .1. **p < .05. ***p < .01.

Notes

1. Poorer White students have been shown to lag behind in some measures of academic achievement, such as lower test score gains in the United Kingdom relative to minority students (Wilson et al., 2009).

2. Family financial resources as measured by socioeconomic status or parental income consistently predict high school completion, whereas the social capital conferred by parents relates to attainment outcomes in varying degrees (Rumberger, 2011). School quality, in particular the social compositions of the student body, was theorized by Coleman et al. (1966) as meaningful inputs to educational achievement and attainment, although existing research offers mixed evidence on their precise contributions (Autor et al., 2016; Cook & Evans, 2000; Fryer & Levitt, 2004; Jencks, 1972).

3. While the High School Longitudinal Study of 2009 contains arguably the most recent high school student cohorts, it is

missing individual-level behavioral attributes such as substance use. Therefore, we do not rely on this data set for our analyses.

4. A substantial literature documents the lower noncognitive skills possessed by GED recipients relative to high school graduates (Heckman et al., 2006; Heckman & Rubinstein, 2001). In excluding GED recipients, we adhere to the reporting conventions of the federal government, which since the 2001 passage of the No Child Left Behind Act has excluded GED recipients from official measures of high school graduation. Similarly, previous studies such as Murnane (2013) and Heckman and LaFontaine (2010) have also excluded GED recipients from calculations of high school graduation rates.

5. We follow the National Center for Education Statistics in conditioning our graduation rate on the set of students that have successfully completed eighth grade and are entering ninth grade. See https://nces.ed.gov/FastFacts/display.asp?id=805

6. The global measure sums up the extent of agreement with the following six items, measured on a Likert-type scale: 1 = *You have a lot of good qualities*; 2 = *You have a lot to be proud of*; 3 = *You like yourself just the way you are*; 4 = *You feel like you are doing everything just about right*; 5 = *You feel socially accepted*; 6 = *You feel loved and wanted*.

7. Students are asked on a scale of 1 to 5 how much they want to go to college as well as their estimated likelihood that they will actually go to college.

8. Children qualify for free lunch under the federally subsidized National School Lunch program if they reside in a household with incomes at or below 130% of the federal poverty level, and they qualify for reduced lunch if household income is between 130% and 185% of the stated poverty level. In North Carolina, children from households that are eligible for receiving Supplemental Nutrition Assistance Program or Cash Assistance are automatically eligible for free meals.

9. Following North Carolina's convention for calculating 4-year graduation rates, we restrict to students who are in ninth grade for the first time.

10. School exit files provide the rationale for leaving, which includes graduation and dropping out, along with other reasons such as transfers to home schooling and private schools. We cross-check graduation and dropout outcomes with reasons given in school exit data. For instance, we code students as having graduated or dropped out if they are marked as such in either the graduation/dropout or exit files.

11. We designate urban schools as those located in cities with a population of at least 250,000, and rural schools as those in locations with fewer than 2,500 people or ZIP codes designated as rural by the Census Bureau.

12. These primarily comprise students who exit from the public school system and therefore can no longer be observed via graduation files. NCERDC exempts students who exit for the following reasons: leaving for another school in the system, leaving to another system within the state, leaving for a different state, death, visiting student status, transfers to within-state private schools and home schools, unconfirmed transfers including transfers to detention centers, and students who do not in fact belong in the cohort.

13. Graduation rates for Add Health and NLSY97 are within 0.5 percentage points of those documented by Heckman and LaFontaine (2010), hereafter HL, for the same data sets. These small differences are likely driven by changes in the sample construction relative to HL. Namely, we use more recent graduation data from both national samples. Our classification of respondents as high school graduates in the NLSY97 uses information through the 2015 survey wave, whereas HL use information through 2008. We also use Wave IV data from Add Health while HL rely on Wave III graduation outcomes.

14. We compare our estimates with official 4-year graduation rates for students entering ninth grade in the 2008 to 2009 through 2010 to 2011 school years from the North Carolina State Board of Education Accountability Services Division. See http://www. ncpublicschools.org/accountability/reporting/cohortgradrate. The graduation rates that we find by gender and race are all within 1.2 percentage points of the values reported by the State Board of Education. The most likely reason for the difference is our exclusion of students with missing data. The relatively lower graduation rate we compute for economically disadvantaged students relative to the state's is due to differences in classification. North Carolina uses a proprietary measure based on a student's FRL status, while we use an indicator for a student ever being FRL eligible during elementary and middle school.

15. One reason for the difference is that the NCERDC includes students who are approximately 10 to 15 years younger. The majority of the Add Health and NLSY97 samples comes from the birth cohorts of 1976–1982 and 1980–1984, respectively. In comparison, North Carolina students were born in the mid-1990s, as they were in ninth grade during 2008–2011. Another source of difference is the definition of high school graduation. We report this measure in national surveys as ever having graduated with a regular diploma by the latest wave, while we report a 4-year adjusted graduation rate in the state administrative data. Note that the disparities we document are consistent with those found in other studies. For example, this is corroborated by evidence showing that students in North Carolina graduate at a higher rate than the nation as a whole (McFarland et al., 2018).

16. We focus on the attainment outcomes of African American students in this study, but these analyses can be applied to Hispanic students in future work. Understanding the factors underlying high school graduation for this group is relevant given that Hispanic students comprise a growing share of school-age children in the United States.

17. In estimates not shown, we find large graduation gaps by socioeconomic status (SES), where advantaged students graduate at a rate between 19.1 and 21.9 percentage points higher than low-income students. Furthermore, the consistency of these gaps across data sets suggests that the divergent attainment records among economically advantaged and disadvantaged students are not driven by one sample, cohort, or time period.

18. Fryer and Levitt (2004) is relatively unique in the literature in finding that the vast majority of the Black-White achievement gap can be accounted for using a set of socioeconomic controls. While they observe this result for kindergarteners, the explanatory power of SES diminishes significantly by the time students reach third grade (Fryer & Levitt, 2006). Murnane et al. (2006) documents that one third of the math achievement gap during kindergarten and third grade is eliminated after accounting for family covariates such as SES background. Clotfelter et al. (2009) find that a vector of socioeconomic factors including FRL eligibility, parental education, gender, age, and district and region type explain

approximately one third of Black-White gaps in math and reading from Grades 3 to 8.

19. Note that we report unweighted graduation rates for the Add Health and NLSY97 samples in the fourth panel. We began Table 1 by weighting our descriptive results in order to show that our samples were consistent with other samples used for similar studies in the literature. As we transition to smaller subsamples of low-income students by race, we subsequently use covariates to probe the source of these gaps. Outside of reporting representative population averages, weighting is not necessarily advised (Solon et al., 2015). One reason is that we do not have a heteroskedasticity argument based on the relative size of White and Black low-income female groups in our sample. Second, we do not need to weight to correct for endogenous sampling. While the probability of sampling individuals in our sample does depend on their race and socioeconomic status, we condition on both in our regressions. As a result, the probability of sampling is exogenous with regard to our set of covariates.

20. We verify the robustness of these results by imputing missing income data and finding gaps that are close in magnitude (Table A1). Sample sizes for Add Health and NLSY97 increase to 1,052 and 1,095, respectively. The larger samples reflect imputed values for household income in approximately 15% to 20% of observations. We impute Add Health income using Census block racial composition, average per capita income, and average share with at least a college degree, and individual racial/ethnicity characteristics. We impute NLSY97 income using the educational attainment of residential mothers, age of biological mother when respondent was born, Census region of residence, residence in MSA (metropolitan statistical areas), urban versus rural area, and individual racial/ethnicity characteristics.

21. These differences in means are statistically significant at levels greater than 1% in both nationally representative data sets.

22. Note that we follow Fryer and Levitt (2004) in not controlling for additional demographic and family characteristics. In their study, Fryer and Levitt controlled for a large set of covariates beyond a core set of basic demographic and socioeconomic variables and found that the approximately 100 additional variables—such as parental education, income, and family size—failed to significantly explain the race gap.

23. We observe the same qualitative results using Add Health data. In Table A3, an initial Black-White graduation advantage of 6.6 percentage points increases to 7.8 percentage points when conditioning on block group characteristics including urbanicity, share of the population by racial group, average per capita income, and proportion of age 25+ year-old adults with at least a college degree. The use of school fixed effects instead of block group characteristics further increases the graduation advantage to 12.7 percentage points.

24. Test scores refer to the Picture Vocabulary Test in Add Health and the Armed Forces Qualifications Test in the NLSY97.

25. In results not shown, standardized test scores among Black females are approximately 0.3 to 0.4 standard deviations lower in North Carolina, while they also have lower GPAs.

26. Heckman et al. (2006) show that the effect of noncognitive skills on the probability of graduating from high school and continuing to college are comparable in magnitude to that of cognitive skills. This literature has also used factor models to demonstrate the role that differences in noncognitive abilities play

in educational attainment across gender and race (Aucejo, 2015; Urzúa, 2008).

27. The present study measures substance use during early adolescence. Some evidence supports a convergence of tobacco, alcohol, and illicit drug use later on during the life course (Pampel, 2008; Watt, 2008).

28. As with the results on cognitive factors, our findings do not change when we impute missing values on family income and socioemotional measures. Table A5 shows a graduation advantage among low-income Black females ranging in magnitude from 5.9 to 6.2 percentage points. Coefficient magnitudes are also similar when including the full set of socioemotional covariates.

29. Suppose the following base and full specifications:

$$Y = \beta^{base} Black + \varepsilon^{base}$$
$$Y = \beta^{full} Black + X\gamma^{full} + \varepsilon^{full},$$

where X is a vector of k additional covariates. The portion of the coefficient explained by new covariates is $\hat{\delta}_{Black} = \hat{\beta}^{base} - \hat{\beta}^{full}$. The contributions of the two channels discussed above are expressed via the following decomposition:

$$\hat{\delta}_{Black} = \sum_k \hat{\Gamma}_{k,Black} \hat{\gamma}_k^{full},$$

where $\hat{\Gamma}_{k,Black}$ is the coefficient on the Black female indicator from an auxiliary regression with the k-th covariate as the dependent variable in the base model. It captures the Black-White difference in a covariate such as substance use after partialing out the effects of explanatory factors in the base regression. $\hat{\gamma}_k^{full}$ corresponds to coefficient on the covariate in the full regression, which measures the extent to which factors like substance abuse are correlated with educational attainment.

30. Previous research document lower educational achievement and attainment stemming from adolescent experiences such as absenteeism and contacts with the juvenile justice system (Goodman, 2014; Hjalmarsson, 2008; Kirk & Sampson, 2013). Our finding that accounting for these factors does not meaningfully affect the racial gap in high school graduation does not preclude their potential causal influence on educational trajectories.

31. The adverse academic effects of cannabis use in particular are hypothesized to operate through two main mechanisms: (1) impaired brain development and cognitive functioning or (2) reduced effort (Marie & Zölitz, 2017; McCaffrey et al., 2010; Pope et al., 2003). For instance, McCaffrey et al. (2010) interpret higher dropout rates as driven by parental and peer influences, while Marie and Zölitz (2017) view evidence as consistent with the cognitive functioning channel.

32. Recent findings underscore the need for a more informed perspective on the social determinants of health and how it may be distinct for White individuals, in particular White women (Astone et al., 2015; Malat et al., 2018).

33. Boys' noncognitive skill development is shown to be more sensitive to the home environment, in particular family disruptions (Bertrand & Pan, 2013). Behavioral problems early on also predict outcomes more for boys than for girls (Owens, 2016). Autor et al. (2019) conduct a comprehensive analysis of the determinants of girls' educational advantage using administrative data from Florida. The authors find this advantage is larger in low-SES

families, as disadvantaged boys' behavioral and academic outcomes are more adversely affected than their sisters'. They attribute the disadvantage gradient in the gender gap to the postnatal environment.

34. Wages for high school graduates were nearly 30% greater than those who completed schooling up to eighth grade, and the analogous premium was 65% for college graduates relative to high school graduates (Goldin & Katz, 2007).

References

Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, *70*(2), 87–107. https://doi.org/10.2307/2673158

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. D. (2011). *Personality psychology and economics* [Technical report]. National Bureau of Economic Research. https://doi.org/10.3386/w16822

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. In O. Ashenfelter, R. Layard, & D. Card (Eds.), *Handbook of labor economics* (pp. 3143–3259). Elsevier. https://doi.org/10.1016/S1573-4463(99)30039-0

Astone, N. M., Martin, S., & Aron, L. (2015). *Death rates for US women ages 15 to 54* [Technical report]. Urban Institute.

Aucejo, E. (2015). *Explaining cross-racial differences in the educational gender gap* (CEP Discussion Paper No 1220). Centre for Economic Performance.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016). School quality and the gender gap in educational achievement. *American Economic Review*, *106*(5), 289–295. https://doi.org/10.1257/aer.p20161074

Autor, D., Figlio, D. N., Karbownik, K., Roth, J., & Wasserman, M. (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal*, *11*(3), 338–381. https://doi.org/10.1257/app.20170571

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, *344*(6186): 843–851. https://doi.org/10.1126/science.1251868

Bertrand, M., & Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, *5*(1), 32–64. https://doi.org/10.1257/app.5.1.32

Borghans, L., Ter Weel, B., & Weinberg, B. (2014). People skills and the labor-market outcomes of underrepresented groups. *Industrial and Labor Relations Review*, *67*(2), 287–334. https://doi.org/10.1177/001979391406700202

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton University Press. https://doi.org/10.1515/9781400831463

Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America: Educational reform and the contradictions of economic life*. Routledge & Kegan Paul.

Bowles, S., Gintis, H., & Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature*, *39*(4), 1137–1176. https://doi.org/10.1257/jel.39.4.1137

Breslau, J., Lane, M., Sampson, N., & Kessler, R. C. (2008). Mental disorders and subsequent educational attainment in a US national sample. *Journal of Psychiatric Research*, *42*(9), 708–716. https://doi.org/10.1016/j.jpsychires.2008.01.016

Breslau, J., Miller, E., Joanie Chung, W.-J., & Schweitzer, J. B. (2011). Childhood and adolescent onset psychiatric disorders, substance use, and failure to graduate high school on time. *Journal of Psychiatric Research*, *45*(3), 295–301. https://doi.org/10.1016/j.jpsychires.2010.06.014

Card, D. (1999). The causal effect of education on earnings. *Review of Educational Research*, *73*(4), 441–449.

Carpenter, C., & Dobkin, C. (2009). The effect of alcohol consumption on mortality: Regression discontinuity evidence from the minimum drinking age. *American Economic Journal: Applied Economics*, *1*(1), 164–182. https://doi.org/10.1257/app.1.1.164

Case, A., & Deaton, A. (2017). Mortality and morbidity in the 21st century. *Brookings Papers on Economic Activity*, *2017*(1), 397–476. https://doi.org/10.1353/eca.2017.0005

Catterall, J. S. (1998). Risk and resilience in student transitions to high school. *American Journal of Education*, *106*(2), 302–333. https://doi.org/10.1086/444184

Chapman, C., Laird, J., Ifill, N., & KewalRamani, A. (2011). *Trends in high school dropout and completion rates in the United States: 1972–2009*. https://nces.ed.gov/pubs2012/2012006.pdf

Chatterji, P. (2006). Illicit drug use and educational attainment. *Health Economics*, *15*(5), 489–511. https://doi.org/10.1002/hec.1085

Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2018). *Race and economic opportunity in the United States: An intergenerational perspective* [Working paper]. National Bureau of Economic Research. https://doi.org/10.3386/w24441

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2009). The academic achievement gap in grades 3 to 8. *Review of Economics and Statistics*, *91*(2), 398–419. https://doi.org/10.1162/rest.91.2.398

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. National Center for Education Statistics.

Cook, M. D., & Evans, W. N. (2000). Families or schools? Explaining the convergence in White and Black academic performance. *Journal of Labor Economics*, *18*(4), 729–754. https://doi.org/10.1086/209975

Crosnoe, R. (2006). The connection between academic failure and adolescent drinking in secondary school. *Sociology of Education*, *79*(1), 44–60. https://doi.org/10.1177/003804070607900103

Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and non-cognitive skill formation. *Journal of Human Resources*, *43*(4), 738–782. https://doi.org/10.3368/jhr.43.4.738

Cunha, F., Heckman, J., & Schennach, S. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, *78*(3), 883–931. https://doi.org/10.3982/ECTA6551

Evensen, M., Lyngstad, T. H., Melkevik, O., & Mykletun, A. (2016). The role of internalizing and externalizing problems in adolescence for adult educational attainment: Evidence from sibling comparisons using data from the Young HUNT Study. *European Sociological Review*, *32*(5), 552–566. https://doi.org/10.1093/esr/jcw001

Farkas, G. (2003). Cognitive skills and noncognitive traits and behaviors in stratification processes. *Annual Review of Sociology*, *29*, 541–562. https://doi.org/10.1146/annurev.soc.29.010202.100023

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, *59*(2), 117–142. https://doi.org/10.3102/00346543059002117

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Fryer, R. G., & Levitt, S. D. (2004). Understanding the Black-White test score gap in the first two years of school. *Review of Economics and Statistics*, *86*(2), 447–464. https://doi.org/10.1162/003465304323031049

Fryer, R. G., & Levitt, S. D. (2006). The Black-White test score gap through third grade. *American Law and Economics Review*, *8*(2), 249–281. https://doi.org/10.1093/aler/ahl003

Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? *Journal of Labor Economics*, *34*(2), 509–543. https://doi.org/10.1086/683668

Gelman, A., & Auerbach, J. (2016). Age-aggregation bias in mortality trends. *Proceedings of the National Academy of Sciences of the U S A*, *113*(7), E816–E817. https://doi.org/10.1073/pnas.1523465113

Goldin, C. D., & Katz, L. F. (2007). Long-run changes in the wage structure: Narrowing, widening, polarizing. *Brookings Papers on Economic Activity*, *2007*(2), 135–165. https://doi.org/10.1353/eca.2008.0000

Goodman, J. (2014). *Flaking out: Student absences and snow days as disruptions of instructional time* [Working Paper No. 20221]. National Bureau of Economic Research. https://doi.org/10.3386/w20221

Grogger, J., & Eide, E. (1995). Changes in college skills and the rise in the college wage premium. *Journal of Human Resources*, *30*(2), 280–310. https://doi.org/10.2307/146120

Hall, W., & Degenhardt, L. (2009). Adverse health effects of non-medical cannabis use. *Lancet*, *374*(9698), 1383–1391. https://doi.org/10.1016/S0140-6736(09)61037-0

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, *46*(3), 607–668. https://doi.org/10.1257/jel.46.3.607

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry*, *46*(3), 289–324. https://doi.org/10.3386/w14064

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*(4), 451–464. https://doi.org/10.1016/j.labeco.2012.05.014

Heckman, J. J., & Krueger, A. B. (2005). *Inequality in America: What role for human capital policies?* MIT Press.

Heckman, J. J., & LaFontaine, P. A. (2010). The American high school graduation rate: Trends and levels. *Review of Economics and Statistics*, *92*(2), 244–262. https://doi.org/10.1162/rest.2010.12366

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review*, *91*(2), 145–149. https://doi.org/10.1257/aer.91.2.145

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411–482. https://doi.org/10.1086/504455

Hjalmarsson, R. (2008). Criminal justice involvement and high school completion. *Journal of Urban Economics*, *63*(2), 613–630. https://doi.org/10.1016/j.jue.2007.04.003

Jackson, M. (2009). Understanding links between adolescent health and educational attainment. *Demography*, *46*(4), 671–694. https://doi.org/10.1353/dem.0.0078

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, *126*(5), 2072–2107. https://doi.org/10.1086/699018

Jencks, C. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. Basic Books.

Jencks, C. (1979). *Who gets ahead?: The determinants of economic success in America*. Basic Books.

Kawaguchi, D. (2004). Peer effects on substance use among American teenagers. *Journal of Population Economics*, *17*(2), 351–367. https://doi.org/10.1007/s00148-003-0158-x

Kessler, R. C., & Saunders, W. B. (1995). Social consequences of psychiatric disorders, I: Educational attainment. *American Journal of Psychiatry*, *152*(7), 1026–1032. https://doi.org/10.1176/ajp.152.7.1026

Keyes, K. M., Vo, T., Wall, M. M., Caetano, R., Suglia, S. F., Martins, S. S., Galea, S., & Hasin, D. (2015). Racial/ethnic differences in use of alcohol, tobacco, and marijuana: Is there a crossover from adolescence to adulthood? *Social Science & Medicine*, *124*, 132–141. https://doi.org/10.1016/j.socscimed.2014.11.035

Kirk, D. S., & Sampson, R. J. (2013). Juvenile arrest and collateral educational damage in the transition to adulthood. *Sociology of Education*, *86*(1), 36–62. https://doi.org/10.1177/0038040712448862

Lundborg, P. (2006). Having the wrong friends? Peer effects in adolescent substance use. *Journal of Health Economics*, *25*(2):214–233. https://doi.org/10.1016/j.jhealeco.2005.02.001

Macleod, J., Oakes, R., Copello, A., Crome, I., Egger, M., Hickman, M., Oppenkowski, T., Stokes-Lampard, H., & Smith, G. D. (2004). Psychological and social sequelae of cannabis and other illicit drug use by young people: A systematic review of longitudinal, general population studies. *Lancet*, *363*(9421), 1579–1588. https://doi.org/10.1016/S0140-6736(04)16200-4

Malat, J., Mayorga-Gallo, S., & Williams, D. R. (2018). The effects of whiteness on the health of whites in the USA. *Social Science & Medicine*, *199*, 148–156. https://doi.org/10.1016/j.socscimed.2017.06.034

Marie, O., & Zölitz, U. (2017). "High" achievers? Cannabis access and academic performance. *Review of Economic Studies*, *84*(3), 1210–1237. https://doi.org/10.1093/restud/rdx020

McCaffrey, D. F., Pacula, R. L., Han, B., & Ellickson, P. (2010). Marijuana use and high school dropout: The influence of unobservables. *Health Economics*, *19*(11), 1281–1299. https://doi.org/10.1002/hec.1561

McDaniel, A., DiPrete, T. A., Buchmann, C., & Shwed, U. (2011). The black gender gap in educational attainment: Historical trends and racial comparisons. *Demography*, *48*(3), 889–914. https://doi.org/10.1007/s13524-011-0037-0

McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., Barmer, A., Cataldi, E., & Mann, F. (2018). The condition of education 2018. *National Center for Education Statistics*. https://nces.ed.gov/pubs2018/2018144.pdf

Mishel, L., & Roy, J. (2006). *Rethinking high school graduation rates and trends*. Economic Policy Institute.

Murnane, R. J. (2013). U.S. high school graduation rates: Patterns and explanations. *Journal of Economic Literature*, *51*(2), 370–422. https://doi.org/10.1257/jel.51.2.370

Murnane, R. J., Willett, J. B., Braatz, M. J., & Duhaldeborde, Y. (2001). Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. *Economics of Education Review*, *20*(4), 311–320. https://doi.org/10.1016/S0272-7757(00)00056-X

Murnane, R. J., Willett, J. B., Bub, K. L., McCartney, K., Hanushek, E., & Maynard, R. (2006). Understanding trends in the black-white achievement gaps during the first years of school [with comments]. *Brookings-Wharton Papers on Urban Affairs*, *2006*, 97–135. https://doi.org/10.1353/urb.2006.0024

Owens, J. (2016). Early childhood behavior problems and the gender gap in educational attainment in the United States. *Sociology of Education*, *89*(3), 236–258. https://doi.org/10.1177/0038040716650926

Paglin, M., & Rufolo, A. M. (1990). Heterogeneous human capital, occupational choice, and male-female earnings differences. *Journal of Labor Economics*, *8*(1), 123–144. https://doi.org/10.1086/298239

Pampel, F. C. (2008). Racial convergence in cigarette use from adolescence to the mid-thirties. *Journal of Health and Social Behavior*, *49*(4), 484–498.

Pope, H. G., Gruber, A. J., Hudson, J. I., Cohane, G., Huestis, M. A., & Yurgelun-Todd, D. (2003). Early-onset cannabis use and cognitive deficits: What is the nature of the association? *Drug and Alcohol Dependence*, *69*(3), 303–310. https://doi.org/10.1016/S0376-8716(02)00334-4

Reardon, S. F., & Robinson, J. P. (2008). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In H. A. Ladd, & E. B. Fiske (Eds.) *Handbook of research in education finance and policy* (pp. 497–516). Lawrence Erlbaum.

Rumberger, R. W. (2011). *Dropping out*. Harvard University Press. https://doi.org/10.4159/harvard.9780674063167

Saunders, J., Davis, L., Williams, T., & Williams, J. H. (2004). Gender differences in self-perceptions and academic outcomes: A study of African American high school students. *Journal of Youth and Adolescence*, *33*(1), 81–90. https://doi.org/10.1023/A:1027390531768

Solon, G., Haider, S. J., & Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources*, *50*(2), 301–316. https://doi.org/10.3368/jhr.50.2.301

Stillwell, R., Sable, J., & Plotts, C. (2011). *Public school graduates and dropouts from the common core of data: School year 2008-09*.

Strand, S. (2014). School effects and ethnic, gender and socio-economic gaps in educational achievement at age 11. *Oxford Review of Education*, *40*(2), 223–245. https://doi.org/10.1080/03054985.2014.891980

Swendsen, J., Burstein, M., Case, B., Conway, K. P., Dierker, L., He, J., & Merikangas, K. R. (2012). Use and abuse of alcohol and illicit drugs in US adolescents: Results of the National Comorbidity Survey–Adolescent Supplement. *Archives of General Psychiatry*, *69*(4), 390–398. https://doi.org/10.1001/archgenpsychiatry.2011.1503

Tamborini, C. R., Kim, C., & Sakamoto, A. (2015). Education and lifetime earnings in the United States. *Demography*, *52*(4), 1383–1407. https://doi.org/10.1007/s13524-015-0407-0

Urzúa, S. (2008). Racial labor market gaps the role of abilities and schooling choices. *Journal of Human Resources*, *43*(4), 919–971. https://doi.org/10.1353/jhr.2008.0011

Watt, T. T. (2008). The race/ethnic age crossover effect in drug use and heavy drinking. *Journal of Ethnicity in Substance Abuse*, *7*(1), 93–114. https://doi.org/10.1080/15332640802083303

Weinberger, C. J. (1999). Mathematical college majors and the gender gap in wages. *Industrial Relations: A Journal of Economy and Society*, *38*(3), 407–413. https://doi.org/10.1111/0019-8676.00134

West, M. R., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F. O., & Gabrieli, J. D. E. (2016). Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis*, *38*(1), 148–170. https://doi.org/10.3102/0162373715597298

Wilson, D., Burgess, S., & Briggs, A. (2009). The dynamics of school attainment of England's ethnic minorities. *Journal of Population Economics*, *24*(2), 681–700. https://doi.org/10.1007/s00148-009-0269-0

Woolf, S. H., Chapman, D. A., Buchanich, J. M., Bobby, K. J., Zimmerman, E. B., & Blackburn, S. M. (2018). Changes in midlife death rates across racial and ethnic groups in the United States: systematic analysis of vital statistics. *BMJ*, *362*, k3096. https://doi.org/10.1136/bmj.k3096

Authors

BRIAN CLARK is an economist in the Antitrust Division of the U.S. Department of Justice. His research focuses broadly on the effect of students' educational choices on labor market outcomes.

YING SHI is an assistant professor of public administration and international affairs at Syracuse University. Her research focuses broadly on the causes and consequences of educational inequality.