Higher Education and the Black-White Earnings Gap*

Xiang Zhou Guanghui Pan

Harvard University University of Oxford

August 17, 2022

Forthcoming, *American Sociological Review*

Abstract

How does higher education shape the Black-White earnings gap? It may help close the gap if Black youth benefit more from attending and completing college than do White youth. On the other hand, Black college-goers are less likely to complete college relative to White students, and this disparity in degree completion helps reproduce racial inequality. In this study, we employ a novel causal decomposition and a debiased machine learning method to isolate, quantify, and explain the equalizing and stratifying roles of college. Analyzing data from the NLSY97, we find that among men, a BA degree has a strong equalizing effect on earnings; yet, at the population level, this equalizing effect is partly offset by unequal likelihoods of BA completion between Black and White students. Moreover, a BA degree narrows the male Black-White earnings gap not by reducing the influence of class background and pre-college academic ability, but by lessening the "unexplained" penalty of being Black in the labor market. To illuminate the policy implications of our findings, we estimate counterfactual earnings gaps under a series of stylized educational interventions. We find that interventions that both boost rates of college attendance and BA completion and close racial disparities in these transitions can substantially reduce the Black-White earnings gap.

*Direct all correspondence to Xiang Zhou, Department of Sociology, Harvard University, 33 Kirkland Street, Cambridge MA 02138; email: xiang_zhou@fas.harvard.edu. Earlier versions of this paper were presented at the 2021 Annual Meeting of the American Sociological Association, the 2021 Spring Meeting of Research Committee 28 of the International Sociological Association, and the 2022 Annual Meeting of the Population Association of America. The authors benefitted from the comments of Jennie Brand, Richard Breen, Christina Ciocca Eller, Aleksei Opacic, Geoffrey Wodtke, the editors and reviewers at *ASR*, and seminar participants at the University of Chicago, University of Wisconsin-Madison, Columbia University, Harvard University, and the Pennsylvania State University.

The disparity between Black and White Americans in economic status is one of the most glaring and unrelenting forms of inequality in the United States. Despite a substantial drop in the mid-20th century, the Black-White gap in earnings has persisted during the post-civil rights era. In 1970, the median annual earnings among Black men were 59% of those among White men; by 2014, this ratio had *worsened* to 50% (Bayer and Charles 2018).[1] Moreover, as shown in a recent study that leverages linked administrative data across two generations, no more than half of the male Black-White gap in individual income can be attributed to racial differences in parental income, parental education, and other markers of socioeconomic background (Chetty et al. 2020b).

To explain the persistence of racial inequality over time, social scientists have proposed an array of accounts that highlight the roles of various micro-, meso-, and macro-level factors, including racial disparities in parental wealth (e.g., Oliver and Shapiro 2006; Conley 2010), family structure and stability (e.g., Moynihan 1965; McLanahan and Sandefur 1994; Bloome 2014), school quality and skill formation (e.g., Jencks and Phillips 1998; Dobbie and Fryer Jr 2011), structural changes in the economy (e.g., Wilson 1987, 1996; Manduca 2018), race- and class-based residential segregation (e.g., Massey and Denton 1993; Reardon and Bischoff 2011; Pattillo 2013), labor market discrimination (e.g., Kirschenman and Neckerman 1991; Bertrand and Mullainathan 2004; Pager et al. 2009), and the racialized penal system (e.g., Pettit and Western 2004; Western 2006). In this article, we focus on the role of higher education — an institution widely perceived as a ticket to economic advancement for disadvantaged youth — in shaping racial inequality. Specifically, we examine whether higher education mitigates, maintains, or magnifies the Black-White earnings gap, and how.

We highlight two mechanisms through which the postsecondary system may affect the Black-White earnings gap: *educational returns* and *educational inequality* (Bloome et al. 2018). First, given the substantial economic returns associated with a college degree (Autor 2014), higher education may serve as an engine of upward mobility for low-income African American youth. Several studies suggest that the economic payoff to a college education for students from disadvantaged backgrounds is as large as, if not larger than, that for their more advantaged peers (e.g., Attewell et al. 2007; Maurin

[1]According to our own calculation based on data from the American Community Survey, this ratio was 0.53 in 2018.

and McNally 2008; Brand and Xie 2010; Zimmerman 2014). For example, Brand and Xie (2010) find that young people who are least likely to receive a bachelor's degree — typically students from minority and low-income backgrounds — appear to benefit the most from it. From this perspective, an expansion in higher education, especially one that induces more African American youth into college, would have the potential to reduce the Black-White earnings gap.

On the other hand, as long recognized by stratification scholars, education is also a vehicle of social reproduction (Blau and Duncan 1967; Boudon 1974; Stevens et al. 2008). Even among those who have "made it to college," the postsecondary system reflects and reinforces preexisting inequalities, as minority and low-income students often attend lower-quality institutions and, partly because of this, graduate at lower rates relative to their more privileged peers (Bowen et al. 2009; Ciocca Eller and DiPrete 2018). Thus, to the extent that the economic payoff to a college education stems mostly from the attainment of a BA degree (rather than merely attending college), unequal rates of degree completion may serve to maintain, if not magnify, racial inequality in earnings. From this perspective, policies that focus exclusively on closing gaps in college enrollment (but not in degree completion) may be an insufficient, or even counterproductive, strategy to combat the Black-White earnings gap.

To date, few studies have considered how educational returns and educational inequality jointly shape racial earnings inequality. In fact, studies that highlight the equalizing potential of higher education typically treat college attendance or completion as an *independent variable* and investigate whether the economic returns to college differ across variously defined subpopulations (e.g., Attewell et al. 2007), whereas studies that foreground the stratifying role of higher education typically regard the completion of a BA degree (henceforth BA completion) as a *dependent variable* and examine how it relates to a student's race, class background, and academic preparation (e.g., Bowen et al. 2009). Consequently, it remains unclear how different forces associated with higher education combine to shape the Black-White earnings gap. This study represents our attempt to address this puzzle.

Integrating data from the National Longitudinal Survey Youth, 1997 (NLSY97), the Department of Education's Integrated Postsecondary Education Data System (IPEDS), and the Opportunity Insights project (Chetty et al. 2020a), we investigate how the effect of attending a four-year college on earnings differs between African Americans and non-Hispanic Whites, and, crucially, the sources of the observed effect heterogeneity. Given the above discussion, we would expect that racial differences in the total effect of college attendance are driven by at least two competing forces: (a) Black students may benefit more from the experience of attending college and from completing a BA degree than do their White peers; (b) given college attendance, Black students are less likely to complete a BA degree relative to their White peers, due to their disadvantages in financial resources, academic preparation, college quality, among other factors.

To understand how these competing forces combine to shape the Black-White earnings gap, we employ a novel causal decomposition that partitions the average effect of attending a four-year college on earnings into four distinct components: (a) the direct effect of college attendance (short of a BA degree) on earnings, (b) the likelihood of BA completion given college attendance, (c) the net effect of BA completion on earnings, and (d) the covariance between BA completion and its net effect on earnings. This decomposition can be understood through Figure 1, where arrows (a), (b), and (c) correspond to the first three components described above, and the last component reflects the interaction effect of arrows (b) and (c). If, as the above discussion suggests, African American youth benefit more from attending and completing college but are less likely to complete a BA degree given attendance compared with White youth, then components (a) and (c) will be equalizing (i.e., stronger for African Americans than for Whites) but component (b) will be stratifying (i.e., weaker for African Americans than for Whites). Thus, such a decomposition allows us to isolate and quantify the equalizing and stratifying roles of college. To make our estimates causally plausible and statistically robust, we adjust for a rich set of individual-, family-, and school-level characteristics that may affect a person's selection into and out of college, employ a debiased machine learning approach to estimate all quantities of interest, and conduct a sensitivity analysis to assess the robustness of our findings to unobserved confounding.

[Figure 1 about here]

In addition to assessing the causal effects of attending and completing a four-year college, we construct a set of *potential* Black-White earnings gaps at different levels of education, including *high school graduate*,[2] *college dropout/stopout*, and *college graduate*. These potential earnings gaps can be interpreted as the levels of inequality that would arise within a random sample of Black and White Americans if their educational status was fixed at a given level (Lundberg 2022). We then examine the extent to which these potential earnings gaps can be explained by racial differences in parental income and pre-college academic ability. By doing so, we illuminate the ways in which higher education attenuates or amplifies Black disadvantage: specifically, whether it modifies Black disadvantage by adjusting the influence of pre-college class and academic backgrounds, or by adjusting the part of inequality that cannot be explained by racial differences in class and academic backgrounds — a part more likely driven by labor market factors such as employer discrimination and job access. Similarly, we construct a set of potential and observed Black-White gaps in BA completion rate and examine the extent to which they can be explained by racial differences in parental income, pre-college academic ability, and college characteristics.

Our empirical analyses yield several key findings. First, we find an equalizing role of higher education among men, but not among women. In particular, the net effect of a BA degree on earnings is much larger for Black men than for White men, but it is similar between Black and White women. At the population level, however, this equalizing effect for men is partly offset by unequal likelihoods of BA completion between Black and White students, leading to a modest and statistically insignificant racial difference in the total effect of college attendance.

Second, because Black men benefit much more from completing college than do White men, the potential earnings gap is substantially lower among college graduates than among high-school graduates and college dropouts/stopouts. This educational gradient in the male Black-White earnings gap is almost entirely due to the educational gradient in the amount of "residual inequality," i.e., the part of inequality that cannot be explained by racial differences in pre-college class and academic

[2]In this study, we focus on the effects of attending and completing a four-year college, and thus classify anyone who had a high school diploma or equivalent but did not attend a four-year college, whether or not the person attended/completed a two-year college, as a "high school graduate." As will be discussed later, our results are similar when those who attended only a two-year college are excluded from the analysis.

backgrounds. In fact, Black-White differences in parental income and pre-college ability translate into a similar amount of earnings inequality at different levels of education. This amount accounts for the bulk of the racial earnings gap among men with a BA degree, but it constitutes only about 30% of the earnings gap among less-educated men. Thus, a college degree narrows the male Black-White earnings gap chiefly by mitigating the "unexplained" penalty of being Black in the labor market, rather than by reducing the influence of class background and pre-college cognitive ability.

Finally, the equalizing effect of a BA degree (for men) and the stratifying force associated with unequal likelihoods of BA completion (for both men and women) bear on educational policies aimed at combating the Black-White earnings gap. They suggest that a blanket expansion in college enrollment is unlikely to significantly reduce the Black-White earnings gap, but an across-the-board increase and/or a leveling in BA attainment rate may help. To illuminate the potential impacts of different policies, we estimate counterfactual Black-White earnings gaps under a series of stylized educational interventions, including race-neutral and race-conscious expansions in college attendance and/or BA completion. Our estimates suggest that interventions that both boost rates of college attendance and BA completion and close racial disparities in these transitions can substantially reduce the Black-White earnings gap.

College as an Equalizer

We characterize college as an equalizer if the causal effect of attending and/or completing college on earnings is greater among African Americans than among Whites. Below, we outline several potential mechanisms that may contribute to such an effect heterogeneity. While it is beyond the scope of this study to test each of the following mechanisms, they highlight several racialized obstacles to economic advancement that higher education, or at least a BA degree, may help African American youth circumvent.

First, a BA degree may help alleviate employer discrimination against Black workers. In signaling models of education (Spence 1973; Weiss 1995), educational attainment is not only a proxy for human capital but also a signal for a worker's expected productivity, as gauged by the average productivity among workers with the same level of education. In a recent refinement of these models, Arcidiacono et al. (2010) argue that a BA degree in fact allows workers to directly reveal their idiosyncratic abilities, not just the average ability of a college graduate, to potential employers. Because college graduates typically include grades, majors, and college(s) attended in their résumés, their ability can be accurately and immediately observed in the labor market, which reduces employers' incentives for statistical discrimination (see also Lang and Manove 2011). By contrast, such information is often lacking among those with only a high school education. As a result, in the "high school labor market," employers have stronger incentives to discriminate, statistically or otherwise, on the basis of race. This is compounded by the fact that a large fraction of low-wage jobs in the contemporary U.S. labor market are located in the retail and service industries, where employers tend to place a heavy emphasis on so-called "soft skills," such as motivation, work ethic, and the ability to interact with co-workers and customers. Compared with White and Hispanic workers, Black workers, especially young Black men, are often perceived as lacking in such skills (Moss and Tilly 1996). As shown in several interview and audit studies (e.g., Kirschenman and Neckerman 1991; Waldinger 1993; Pager et al. 2009), racial projections with regard to soft skills constitute a formidable barrier to employment for less-educated Black men. Thus, to the extent that employer incentives to use racial cues are stronger in the hiring and promotion of non-college-educated workers, a BA degree should narrow the Black-White earnings gap through a reduction in labor market discrimination.

To be sure, the above arguments do not imply an absence of racial discrimination against highly-educated African Americans. Using data from the National Survey of College Graduates, Black et al. (2006) find that only a small portion of the Black-White wage gap among college graduates can be explained by racial differences in premarket factors such as college major and type of degree. In an audit study, Gaddis (2015) provides direct evidence of discrimination by showing that even a BA degree from an elite university does not equalize callback rates between Black and White job applicants, and, in fact, "black candidates only do as well as white candidates from less selective universities" (p. 1451). Nonetheless, these studies do not contradict the possibility that labor market discrimination may be reduced, albeit not eliminated, among college graduates.

Higher education may also narrow the Black-White earnings gap through its heterogeneous effects on neighborhood attainment. It is well documented that African Americans are far more likely than Whites to reside in economically distressed communities with few employment opportunities (Wilson 1987, 1996; Massey and Denton 1993), limited access to and return on job referral networks (Mouw 2002; Royster 2003; Smith 2005, 2007; Pedulla and Pager 2019), and low levels of social organization (Sampson and Wilson 1995; Sampson 2012). Moreover, Black children who grow up in poor neighborhoods are more likely to stay in poor neighborhoods as adults than similarly situated White children (Sharkey 2008, 2013). The economic and social isolation of poor Black neighborhoods is long considered a contributor to the employment gap between Black and White Americans (e.g., Kain 1968; Jencks and Mayer 1990; Holzer 1991; Ihlanfeldt and Sjoquist 1998; Mouw 2000; Hellerstein et al. 2008). Yet, attending and completing college often entails relocation and may be a particularly important channel for Black youth to move out of disadvantaged neighborhoods. In fact, several studies suggest an equalizing effect of higher education on neighborhood poverty (South and Crowder 1997; Adelman et al. 2001; Crowder and South 2005; Swisher et al. 2013). For example, analyzing data from the National Longitudinal Study of Adolescent Health, Swisher et al. (2013) find that for young adults in the early 2000s, both college attendance and completion of a four-year degree are "associated with decreases in neighborhood poverty, with blacks receiving a stronger return from educational attainments than whites" (p.1399). In this regard, given the influence of neighborhood poverty on job access, job information networks, and social norms of employment, higher education should help attenuate the Black-White gap in employment and earnings.

In addition, by facilitating neighborhood mobility and a strong attachment to school and work, higher education will likely reduce young adults' exposure to and involvement in illegal activities, thus lowering the risk of arrest and incarceration (Lochner and Moretti 2004). As shown in Pettit and Western (2004), the risk of incarceration among young men is highly stratified by race, education, and, importantly, their interaction. While the likelihood of incarceration does not differ markedly between college-educated Black and White men, it is exceedingly high among Black men without a college education. This interaction is consequential given the large and far-reaching impact of incarceration on a person's life chances (Western 2002, 2006; Wakefield and Uggen 2010; Reich and Prins 2020). A spell of prison time not only precludes opportunities for regular employment in the short run, but also undermines the employment prospects of ex-convicts after their release, due to the stigma of crime, the erosion of human capital, and the disruption of social and family ties, among other factors. Thus, considering the interaction effect of race and education on the risk of incarceration and the deleterious effects of incarceration on employment, higher education may alleviate earnings inequality among men by narrowing the Black-White gap in the risk of incarceration.

Empirical Work on Educational Variations in Racial Earnings Inequality

A number of empirical studies have documented a negative association between the level of education and the Black-White earnings gap among men (Neal and Johnson 1996; Johnson and Neal 1998; Bjerk 2007; Arcidiacono et al. 2010; Lang and Manove 2011; Sakamoto et al. 2018). Using data from the National Longitudinal Survey of Youth, 1979 cohort (NLSY79), Johnson and Neal (1998) report that in terms of both wages and working hours, the college premium is greater among Black men than among White men. Relatedly, Bjerk (2007) finds that after adjusting for cognitive ability, the Black-White wage gap is more pronounced in blue-collar occupations than in white-collar occupations. More recently, using data from the U.S. Social Security Administration linked to the Survey of Income and Program Participation, Sakamoto et al. (2018) find that in terms of lifetime earnings, Black men are more disadvantaged at lower levels of education, and that this educational gradient in the Black-White earnings gap cannot be explained by demographic characteristics, work disability, or measures of academic achievement (see also Cheng et al. 2019).

Empirical evidence is relatively scanty about educational variations in racial earnings inequality among women. Neal (2004) finds that the Black-White wage gap among working women is somewhat smaller among college graduates than among those with lower levels of education. Yet, as the author points out, the wage gap among working women suffers from selection bias due to racial differences in patterns of female labor supply: whereas the rate of overall labor force participation is similar between

Black and White women, Black women who do not work are disproportionately less educated than their White counterparts, leading to an underestimate of the Black-White gap in potential wages. In this study, we do not restrict our analysis to workers. Given the gender and racial differences in patterns of labor supply, we conduct all of our analyses separately for men and for women and discuss gender differences when they appear. More importantly, for both men and women, we provide a more systematic assessment of the causal effects of college on racial earnings inequality by isolating the effects of college attendance and BA completion, by adjusting for a rich set of pre-college and postsecondary characteristics that may shape a person's selection into and out of college, and by employing a debiased machine learning approach that makes efficient use of such high-dimensional data.

College as a Stratifier

Alongside its equalizing roles discussed above, the postsecondary system also maintains and reproduces racial inequality. The stratifying role of higher education is reflected in the fact that given college attendance, Black students are much less likely to complete a BA degree relative to their White peers. Among those who started at a four-year college in 2010, for example, 64% of White students graduated with a BA degree within six years, compared with only 40% of Black students (Jeffrey 2020; see also Snyder et al. 2019). The Black-White disparity in college graduation rates has barely changed over the past several decades (Voss et al. 2022), and, as documented by previous research, it is shaped by a range of factors, including Black students' disadvantages in financial resources, academic preparation, college quality, as well as the psychologically harmful consequences of negative racial stereotypes that Black students regularly face in academic life.

First, the Black-White gap in BA completion partly stems from racial disparities in financial resources. It is well documented that the probability of college graduation is highly stratified by family income. In the NLSY97 cohort, for example, college-goers in the top quartile of the family income distribution are more than twice as likely to graduate by age 25 relative to those in the bottom quartile

(Bailey and Dynarski 2011). Moreover, only a small portion of the gap in college graduation rates between high- and low-income students can be explained by their differences in academic preparation and university attended (Bowen et al. 2009, pp. 40-41), which suggests that financial resources play a more direct role in shaping a student's persistence in college than one might expect (Dwyer et al. 2012). As shown in a number of qualitative studies (e.g., Armstrong and Hamilton 2013; Lee 2016; Stuber 2009, 2011; Jack 2019), students of different economic backgrounds who attend even the same college tend to find themselves on divergent trajectories of campus life. With abundant financial resources, students from more affluent families are often freed from the need to work and thus can fully engage in academic and extracurricular activities, both of which facilitate persistence (Tinto 1994; Braxton et al. 1997). By contrast, students from low-income backgrounds often have to juggle school, work, and family responsibilities, which hurts academic performance and elevates the risk of dropping out. Given that Black college-goers disproportionately come from low-income families, income-based inequality in college persistence has likely contributed to the Black-White gap in BA completion.

The Black disadvantage in BA completion is also a result of racial differences in academic preparation. Measures of pre-college academic achievement, such as high school GPA and test scores, are strongly predictive of graduation. For example, Bowen et al. (2009, p. 115) find that "one standard deviation in high school grades is associated with increases in graduation rates of 10-20 percentage points," a relationship that is statistically significant in all but one of the 52 public universities in their study. Given the well-documented disparity between Black and White students in pre-college academic performance (e.g., Jencks and Phillips 1998; Neal 2006), we would expect academic preparation to play a significant role in producing the Black-White gap in graduation rates. A recent study by Ciocca Eller and DiPrete (2018) lends credence to this view. Applying a regression-based decomposition to data from the Education Longitudinal Study of 2002, the authors find that among various types of pre-college resources, skills, and experiences, academic performance is the strongest contributor to the Black-White disparity in college dropout, explaining nearly half of the total gap.

The Black-White gap in BA completion may also result from racial disparities in college quality. Culling data from several longitudinal studies conducted by the National Center for Educational Statistics, Reardon et al. (2012, p. 2) report that "the probability of enrolling in a highly selective college is five times greater for White students than for Black students," and, even after adjusting for family income, "white students are two to three times as likely as black students to gain admission to highly selective colleges." These disparities are largely driven by race and class differentials in academic preparation. In fact, relative to White students, Black students with comparable class backgrounds and academic qualifications are more likely to attend the nation's most selective colleges and universities, thanks in part to affirmative action policies practiced by those institutions (Grodsky 2007; Ciocca Eller and DiPrete 2018; Conwell and Quadlin 2021). On the other hand, however, it has been found that relative to White students, Black students are also more likely to "undermatch," i.e., to enroll in a college that is less selective than the kind of colleges they would likely have been accepted to given their academic records (e.g., Bowen et al. 2009). Regardless of its sources, the Black-White inequality in college selectivity is consequential because, as shown repeatedly in previous scholarship, attending a more selective college has an independent and sizable effect on the likelihood of BA completion, and this effect appears even greater for Black students than for their White peers (Alon and Tienda 2005; Small and Winship 2007; Melguizo 2008; Bowen et al. 2009). Thus, racial disparities in college selectivity, and in college quality more generally, may have widened the Black-White gap in graduation rates.

In addition to racial differences in financial resources, academic preparation, and college quality, previous scholarship has also highlighted the role of psychological processes in the academic achievement of Black students. The theory of stereotype threat (Steele 1988, 1992; Steele and Aronson 1995), in particular, holds that Black students underperform academically partly because of an unconscious fear of confirming negative societal stereotypes about the mental ability of African Americans. The prospects of being stereotyped, and of failing to disconfirm such negative stereotypes, constitute a psychological threat that directly undermines test performance. In the long term, to reduce anxiety and stress, Black students may disidentify with academic achievement as a metric of self-worth, and thus disengage from schoolwork. The predictions of stereotype threat theory pertaining to test performance have been confirmed in a series of laboratory studies (Steele and Aronson 1995). Furthermore, analyzing survey data from a large sample of students in 28 selective colleges and universities, Massey, Charles, and colleagues (2003) find that Black and Hispanic students who doubted their abilities and were self-conscious about the views of teachers earned significantly lower grades and failed courses much more frequently relative to other minority students. In a follow-up study, Charles et al. (2009) corroborate this finding by showing that a sizable portion of the Black-White GPA gap in these colleges can be explained by indicators of stereotype threat. Thus, given the impact of college grades on the risk of dropping out, psychological factors associated with racial stereotypes may have contributed to the Black-White gap in graduation rates.

Analytical Strategy

To elucidate how the equalizing and stratifying roles of higher education combine to shape the Black-White earnings gap, we employ a novel framework, which draws upon the language and logic of causal mediation analysis, for studying the effects of higher education on earnings. By treating BA completion as a mediator that transmits the effect of college attendance on earnings, it partitions the total effect of attending a four-year college into a direct effect of college attendance (short of a BA degree) and an indirect effect via a BA degree. The latter component is sometimes referred to as the "continuation value" of college attendance (Heckman et al. 2018), and it is governed by a person's likelihood of BA completion given college attendance as well as the net effect of BA completion on earnings. Specifically, for individual i, let A_i denote a binary indicator of attending a four-year college, M_i a binary indicator of BA completion, and Y_i labor market earnings. In addition, using the potential outcomes notation, let $M_i(a)$ denote individual i's potential status of BA completion if her college attendance status was set to a, and let $Y_i(a, m)$ denote individual i's potential earnings if her college attendance status was set to a and BA completion status set to m. The total effect (TE) of college attendance on earnings can be written as

$$
\begin{aligned}
\mathrm{TE}_i &= Y_i\big(1, M_i(1)\big) - Y_i(0, M_i(0)) \\
&= Y_i\big(1, M_i(1)\big) - Y_i(0,0) \quad \text{(because } M_i(0) = 0\text{)} \\
&= \underbrace{Y_i(1,0) - Y_i(0,0)}_{\substack{\text{direct effect of} \\ \text{college attendance}}} + \underbrace{M_i(1)\,\underbrace{\big(Y_i(1,1) - Y_i(1,0)\big)}_{\substack{\text{net effect of} \\ \text{BA completion}}}}_{\substack{\text{indirect effect via} \\ \text{BA completion}}}.
\end{aligned} \tag{1}
$$

Thus, for individual i, the total effect of college attendance is governed by three components: the direct effect of college attendance ($Y_i(1,0) - Y_i(0,0)$), whether the person would complete a BA degree given college attendance ($M_i(1)$), and the net effect of BA completion ($Y_i(1,1) - Y_i(1,0)$). The product of the latter two components constitutes the indirect effect of college via BA completion.

To see how the direct and indirect effects of college differ by race, and how racial differences differ by gender, we focus on the conditional mean of TE_i in each subpopulation defined by gender and race, denoted by G. In light of equation (1) and the fact that $\mathbb{E}[UV] = \mathbb{E}[U]\mathbb{E}[V] + \mathrm{cov}[U,V]$ for any random variables U and V, this conditional mean can be decomposed into four components:

$$
\begin{aligned}
\underbrace{\mathbb{E}[\mathrm{TE}_i|G_i = g]}_{=:\Delta^g_{\text{tot}}} =& \mathbb{E}[Y_i(1,0) - Y_i(0,0)|G_i = g] + \mathbb{E}[M_i(1)\big(Y_i(1,1) - Y_i(1,0)\big)|G_i = g] \\
=& \underbrace{\mathbb{E}[Y_i(1,0) - Y_i(0,0)|G_i = g]}_{=:\Delta^g_{\text{att}}} + \underbrace{\mathbb{E}[M_i(1)|G_i = g]}_{=:\pi^g_{\text{comp}}} \cdot \underbrace{\mathbb{E}[Y_i(1,1) - Y_i(1,0)|G_i = g]}_{=:\Delta^g_{\text{comp}}} \\
&+ \underbrace{\mathrm{Cov}[M_i(1), Y_i(1,1) - Y_i(1,0)|G_i = g]}_{=:\Delta^g_{\text{cov}}} \\
=& \Delta^g_{\text{att}} + \underbrace{\pi^g_{\text{comp}}\Delta^g_{\text{comp}} + \Delta^g_{\text{cov}}}_{=:\Delta^g_{\text{ind}}},
\end{aligned} \tag{2}
$$

Thus, for group g, Δ^g_{tot} reflects the average total effect of college on earnings, Δ^g_{att} reflects the average direct effect of college attendance on earnings, π^g_{comp} reflects the probability of BA completion given college attendance, Δ^g_{comp} reflects the average net effect of BA completion on earnings, and

Δ^g_{cov} reflects the covariance between BA completion and its net effect on earnings. The last three components together ($\pi^g_{\text{comp}}\Delta^g_{\text{comp}} + \Delta^g_{\text{cov}}$) constitute the average indirect effect via BA completion (Δ^g_{ind}).

The above decomposition enables us to isolate and quantify the equalizing and stratifying roles of higher education discussed previously. Specifically, our preceding arguments suggest that a BA degree may play an equalizing role through a reduction in employer discrimination, and a college education in general may play an equalizing role through its heterogeneous effects on neighborhood disadvantage and incarceration risk. In this regard, both Δ^g_{att} and Δ^g_{comp} may be greater for Blacks than for Whites. On the other hand, given the Black-White disparity in the likelihood of college graduation (as a result of racial differences in financial resources, academic preparation, college quality, and psychological processes), π^g_{comp} is expected to be larger for Whites than for Blacks. Table 1 sums up the implications of the equalizing and stratifying roles of college discussed earlier for different components of equation (2).

[Table 1 about here]

The covariance component Δ^g_{cov} characterizes the pattern and degree of selection into college completion. Specifically, $\Delta^g_{\text{cov}} > 0$ if those who would benefit more from a BA degree (higher values of $Y_i(1,1) - Y_i(1,0)$) are more likely to complete college given attendance ("positive selection"), and $\Delta^g_{\text{cov}} < 0$ if those who would benefit less from a BA degree (lower values of $Y_i(1,1) - Y_i(1,0)$) are more likely to complete college given attendance ("negative selection"). The comparative advantage argument in labor economics (Willis and Rosen 1979; Carneiro et al. 2011) suggests that college students may possess knowledge about their idiosyncratic payoffs to a BA degree and act on it, leading to a pattern of positive selection ($\Delta^g_{\text{cov}} > 0$). On the other hand, a pattern of negative selection ($\Delta^g_{\text{cov}} < 0$) may arise if those who would benefit more from a BA degree face stronger structural barriers that lower their probability of completion (Brand and Xie 2010). The group-specific decomposition (2) allows us to see how the pattern and degree of selection differ by gender and race. For example, if positive selection is present and if it is stronger among White students than among Black students (e.g., due to unequal access to information), then the covariance component

Δ^g_{cov} may also be greater for Whites than for Blacks, contributing to the stratifying role of college.

[Figure 2 about here]

Capitalizing on a rich set of individual-, family-, and school-level characteristics that may affect a person's selection into and out of college, all components in equation (2) can be identified under the assumption of sequential ignorability (Robins 1997). Specifically, if we use X to denote a set of background characteristics that may confound the causal effects of college attendance and BA completion on earnings, and Z to denote a set of postsecondary characteristics that may confound the causal effect of BA completion on earnings, such as college quality and college GPA, the sequential ignorability assumption states that (a) conditional on background characteristics X, no unobserved confounding exists for the effect of college attendance on college completion status and earnings; (b) among college-goers, conditional on background characteristics X and postsecondary characteristics Z, no unobserved confounding exists for the effect of BA completion on earnings. The sequential ignorability assumption is satisfied in Figure 2, which contains a directed acyclic graph (DAG) visualizing a set of hypothesized causal relationships between the variables defined previously.[3] Note that the postsecondary characteristics Z are defined only among college-goers ($A = 1$). Potential values of these variables for non-college-goers are not required to identify and estimate our causal effects of interest.

To be sure, our assumption that the postsecondary characteristics Z are causally intermediate between attendance and completion is a simplification of real-world processes governing students' college-going behavior. For example, a student's decision to attend a four-year college may be affected by *potential* postsecondary characteristics, such as the availability of a scholarship that can defray

[3] In our framework, the postsecondary characteristics Z play two roles. On the one hand, they mediate the effect of college attendance on earnings, as reflected in the paths $A \rightarrow Z \rightarrow Y$ and $A \rightarrow Z \rightarrow M \rightarrow Y$. The first path ($A \rightarrow Z \rightarrow Y$) is captured in the direct effect of college attendance (Δ^g_{att}), and the second path ($A \rightarrow Z \rightarrow M \rightarrow Y$) in its indirect effect (Δ^g_{ind}). On the other hand, the postsecondary characteristics are confounders of the relationship between BA completion and earnings, which means that the net effect of BA completion Δ^g_{comp} reflects the effect of a BA degree *above and beyond* the effects of both college attendance per se and the college experience captured in Z, hence the name "net effect."

the costs of college. If such a scholarship also affects the student's BA completion status and/or earnings, we view it as an unobserved *pre-college* confounder. To assess the robustness of our results to this and other types of unobserved confounding, we conduct a sensitivity analysis that investigates the direction and magnitude of potential bias when either condition (a) or (b) breaks down (see Supplementary Material H for more details).

Under sequential ignorability, both the total effect of college and its direct and indirect components shown in equation (2) are nonparametrically identified, i.e., expressed in terms of observed data without any functional form assumption. The identification formulas of these quantities are given in Supplementary Material A. A variety of methods, such as g-computation (Robins 1997) and marginal structural models (MSMs; Robins et al. 2000), could be used to evaluate these formulas. In this study, we estimate all quantities of interest using a debiased machine learning (DML) approach (Chernozhukov et al. 2018; Semenova and Chernozhukov 2021). In this approach, for each of our target parameters (i.e., the components in equation 2), we construct a so-called "Neyman-orthogonal signal," which is a function of observed data for each unit. For example, our Neyman-orthogonal signal of the average potential earnings under non-college-attendance ($\mathbb{E}[Y(0,0)]$) for individual i is

$$Y_i^*(0,0) = \mathbb{E}[Y_i|X_i, A_i = 0] + \frac{1 - A_i}{1 - \Pr[A_i = 1|X_i]}\big(Y_i - \mathbb{E}[Y_i|X_i, A_i = 0]\big).$$

Such signals satisfy several properties. First, their (conditional) expectations are equal to those of the corresponding potential outcomes; for example, $\mathbb{E}[Y_i^*(0,0)|G_i = g] = \mathbb{E}[Y_i(0,0)|G_i = g]$. Therefore, we can estimate the latter by constructing empirical estimates of $Y_i^*(0,0)$ and then averaging them among members of group g. Second, such estimates are robust in the sense that they remain consistent even if some of the models involved in estimating the Neyman-orthogonal signals are misspecified. For example, estimating $Y_i^*(0,0)$ involves estimating two "nuisance functions"[4]: the conditional mean of earnings ($\mathbb{E}[Y|x,a]$) and the conditional probability of attending

[4]A nuisance function is a function that is not of our primary interest but needed to estimate our target parameters. For example, if we use a propensity-score-based method (e.g., propensity score

college ($\Pr[A = 1|x]$). Yet, our estimate of $\mathbb{E}[Y_i(0,0)|G_i = g]$ remains consistent even if one of the two nuisance functions is misspecified. Finally, when flexible machine learning methods are used to estimate the nuisance functions, our estimates of the target parameters are not only robust to model misspecification but also efficient (i.e., having relatively small standard errors) (see Zhou [forthcoming] for a more in-depth discussion of this point). To avoid potential overfitting of the nuisance functions by machine learning methods, Chernozhukov et al. (2018) introduce a procedure called "cross-fitting," which involves the use of different subsamples to estimate the nuisance functions and the target parameters (see Supplementary Material A for more details). The term "debiased machine learning," in a nutshell, refers to the combined use of Neyman-orthogonal signals, machine learning estimates of nuisance functions, and cross-fitting. For a systematic and accessible introduction to these concepts, see Kreif and DiazOrdaz (2019).

DML is particularly attractive in our context, in which the rich sets of background characteristics (X) and postsecondary characteristics (Z) (see the next section) make it unrealistic for us to correctly specify parametric models for college attendance and college completion, which would be required to justify conventional methods such as MSMs. By leveraging machine learning methods to fit the models for college attendance, college completion, and earnings, the DML approach is highly robust to model misspecification. In this study, we fit each of the requisite models using a super learner (van der Laan et al. 2007)[5] composed of Lasso and random forests (Hastie et al. 2009; see Athey and Imbens [2019] or Molina and Garip [2019] for an introduction to various machine learning methods from a social science perspective). Meanwhile, through the use of Neyman-orthogonal signals and cross-fitting, the DML estimators avoid the regularization and overfitting biases that often afflict machine learning estimators of statistical parameters (see Chernozhukov et al. [2018] for a demonstration of these biases). Finally, as we will see below, our estimates of the Neyman-orthogonal signals can serve as dependent variables in a variety of regression models, allowing us to better

matching) to estimate the average causal effect of college attendance on earnings, the propensity score model for college attendance is a nuisance function.

[5]A super learner is a weighted average of different machine learning methods designed to minimize prediction error. The algorithm is implemented in the R package SuperLearner (Polley and van der Laan 2017).

understand the contributing factors to racial differences in college returns and in BA completion. The implementation and rationale of the DML approach in our context are described in greater detail in Supplementary Material A.

Besides assessing how the causal effects of college attendance and BA completion differ by race, we construct a set of Black-White gaps in *potential earnings*, i.e., $\mathbb{E}[Y_i(a,m)|\text{Black men}]$-$\mathbb{E}[Y_i(a,m)|\text{White men}]$ and $\mathbb{E}[Y_i(a,m)|\text{Black women}]$-$\mathbb{E}[Y_i(a,m)|\text{White women}]$ for different values of a and m. Such quantities have been called "gap-closing estimands" (Lundberg 2022; see also Jackson and VanderWeele 2018), which can be interpreted as earnings gaps that would arise within a random sample of Blacks and Whites if their educational status was fixed at a given level.[6] Unlike the observed Black-White earnings gaps conditional on educational status, these quantities are adjusted for selection processes, thus reflecting the causal effects of higher education on earnings inequality.

We then examine the sources of racial differences in college returns (if any) and in completion rates. First, we assess the extent to which the potential earnings gaps can be (statistically) explained by Black-White differences in family economic background and pre-college academic ability, and how the explanatory power of these factors varies by education. In other words, we evaluate quantities of the form $\mathbb{E}[Y_i(a,m)|\text{Black men}, W]$-$\mathbb{E}[Y_i(a,m)|\text{White men}, W]$ to see how much of the potential earnings gaps can be attributed to factor W. Specifically, we fit and compare three regression models for the (estimated) Neyman-orthogonal signal of each of the potential outcomes (i.e., $Y_i(0,0)$, $Y_i(1,0)$, and $Y_i(1,1)$): a model controlling only for race, a model controlling for race and parental income, and a model that controls additionally for pre-college academic ability, which is measured by the respondent's percentile score on the Armed Services Vocational Aptitude Battery (ASVAB).[7] This

[6]These hypothetical earnings gaps may be referred to as "controlled disparities," akin to the concept of "controlled mobility" introduced in Zhou (2019). Following Lundberg (2022), we interpret these quantities *locally*, i.e., as gaps that would arise within a random sample of Blacks and Whites if their educational status was fixed at a given level. Alternatively, these quantities could be interpreted *globally*, i.e., as gaps that would arise if the educational status of all Black and White youth in the population was fixed at a given level. Compared with the global interpretation, the local interpretation is more credible as it is relatively immune to potential violations of the stable unit treatment value assumption (SUTVA; Rubin 1986). See Lundberg (2022) for more discussion.

[7]In this analysis, we use linear models with only the main effects of the covariates. These linear models are not assumed to characterize the true conditional means of the potential outcomes; instead,

procedure is justified by the fact that the conditional means of the signals are equal to the conditional means of the corresponding potential outcomes, as noted earlier. Second, we assess the degree to which the Black-White gap in BA completion can be explained by parental income, academic preparation, and college quality. Toward this goal, we fit and compare three regression models for the (estimated) Neyman-orthogonal signal of potential BA completion status given college attendance (i.e., $M_i(1)$): a model controlling only for race, a model controlling for race and parental income, and a model that controls additionally for the ASVAB score. In addition, to assess whether college quality plays an independent role in shaping the Black-White gap in graduation rates, we fit a series of additional models for $M_i(1)$ only among college-goers, controlling successively for race, parental income, the ASVAB score, and measures of college quality.

Data and Measures

The primary data source for this study is the National Longitudinal Survey of Youth, 1997 cohort (NLSY97). The NLSY97 began with a nationally representative sample of 8,984 men and women at ages 12-17 in 1997. These individuals were interviewed annually through 2011 and biennially thereafter. As will be discussed later, we also leverage the NLSY Geocode data and data from the IPEDS and the Opportunity Insights project (Chetty et al. 2020a) to construct a set of college-level characteristics. We limit our analytical sample to White, Black, and Hispanic respondents who had completed at least a high-school diploma or GED by age 22 ($n = 7,117$) and who had valid earnings information at ages 30-33 ($n = 6,126$). While our focus is on the Black-White earnings gap, including Hispanic respondents in our analyses offers a comparative lens for us to interpret the degree and patterns of Black disadvantage.

Previous studies on the economic returns to college have often used data from the NLSY79 (e.g., Brand and Xie 2010; Carneiro et al. 2011). Compared with the NLSY79, the NLSY97 traces the educational and labor market experience of a much younger cohort, making findings from this study

they are a data summary device that helps inform the degree to which the potential earnings gaps can be statistically explained by racial differences in family economic background and pre-college ability.

more pertinent to current and future cohorts of American youth. However, because members of the NLSY97 cohort are still relatively young, we can evaluate the economic payoff to higher education only up to their early thirties. Some previous research suggests that the Black-White earnings gap widens over the life course, especially among the highly educated (Tomaskovic-Devey et al. 2005). To explore how our findings vary across cohorts and over the life course, we conduct a supplementary analysis with data from the NLSY79 cohort, drawing on respondents' earnings measured at different ages. Results from this analysis are reported and discussed in Supplementary Material F.

We construct five sets of variables, each corresponding to a node in Figure 2: college attendance (A), BA completion (M), earnings (Y), pre-college characteristics (X), and postsecondary characteristics (Z). Specifically, college attendance (A) denotes whether the respondent had attended a four-year college by age 22, and BA completion (M) denotes whether the respondent had received a BA degree by age 29. A respondent is coded as a *college-goer* (i.e., $A = 1$) if she had either attended a four-year college by age 22 or received a BA degree by age 29, and as a *high school graduate* otherwise (i.e., $A = 0$). Among college-goers, a respondent is coded as a *college graduate* (i.e., $M = 1$) if she had received a BA degree by age 29, and as a *college dropout/stopout* (i.e., $M = 0$) otherwise.[8] By our definition of college attendance, those who had not attended a four-year college by age 22 or received a BA degree by age 29 are coded as high school graduates, whether or not they ever attended a two-year college or attended a four-year college only after age 22. To assess the robustness of our findings, we have conducted parallel analyses using a range of alternative definitions of college attendance and BA completion. The results, as reported in Supplementary Materials C and D, are similar across different specifications.

Earnings are defined as the average of annual earnings, which include incomes from wages, salaries, farms, and other businesses, at ages 30-33 (inflation-adjusted to 2019 dollars). To account for the right skewness of earnings, most existing studies on the Black-White earnings gap have used log earnings as the dependent variable and excluded individuals with zero earnings (e.g., Johnson

[8] College-goers include both those who started college at a four-year school and transfer students from two-year schools. In our sample, 28.8% of college-goers had previously attended a two-year college.

and Neal 1998). Because employment rates are highly unequal between Black and White men (due to racial disparities in unemployment, labor force nonparticipation, and incarceration), excluding individuals with zero earnings will likely distort our results on the degree and patterns of racial inequality (Western and Pettit 2005; Bayer and Charles 2018). Instead, we include all individuals in our analyses but add 1,000 (in 2019 dollars) to the respondent's average annual earnings before taking the log transformation. To assess the sensitivity of our findings to this earnings measure, we have conducted a series of parallel analyses using alternative adjustments for the log transformation as well as the percentile rank transformation. As shown in Supplementary Material E, the degree of the Black-White earnings gap in terms of log points varies according to the constant we add to earnings before taking the log transformation, especially among less-educated men. Nonetheless, our findings about the equalizing role of BA completion are highly consistent across alternative measures of earnings.

To adjust for selection processes that may confound the causal effects of college attendance and BA completion on earnings (i.e., the A-Y and M-Y relationships), we include a broad array of background characteristics (X) in our models for college attendance, BA completion, and earnings. They include basic demographic variables (gender, race, ethnicity, age in 1997), socioeconomic background (parental education, parental income, parental assets, co-residence with both biological parents, presence of a paternal figure, rural residence, southern residence), ability and behavior (percentile score on the ASVAB test, high school GPA, an index of substance use [ranging from 0 to 3], an index of delinquency [ranging from 0 to 10], whether the respondent had any children by age 18), and peer and school-level characteristics (college expectation among peers and three dummy variables denoting whether the respondent ever had property stolen at school, was ever threatened at school, and was ever in a fight at school). In particular, parental education is measured using mother's years of schooling; when mother's years of schooling is unavailable, it is measured using father's years of schooling. Parental income is measured as the average annual parental income from 1997 to 2001. Both parental income and parental assets are inflation-adjusted to 2019 dollars.

To adjust for selection processes that may confound the causal effect of BA completion on earnings among college-goers (i.e., the M-Y relationship), we include a battery of postsecondary characteristics (Z), in addition to the background characteristics X, in our models for BA completion and earnings. They include college type, college quality, field of study, college GPA, and the amounts of student loans. In each survey wave of the NLSY97, respondents were asked to report the names of the colleges in which they were currently or most recently enrolled (if any). Since many respondents attended more than one college, we focus on the college in which the respondent had been enrolled for the longest time by age 29. College type is a trichotomous variable denoting whether the college is a public institution, a private not-for-profit institution, or a for-profit institution. We employ a multi-dimensional measure of college quality that reflects not only admission selectivity but also graduation rate and the college's record of helping low-income students move up the economic ladder. To gauge college selectivity, we use three dummy variables to denote whether the college is one of the "most competitive," "highly competitive," and "very competitive" colleges in Barron's Profile of American Colleges 2000. To measure graduation rate, we use the percentage of students graduating within six years measured in 2002, which is available from the IPEDS database. In addition to college selectivity and graduation rate, we extract from the database of the Opportunity Insights project a measure of "upward mobility rate," i.e., the percentage of students who reach the top quintile of the income distribution among those with parents in the bottom quintile of the income distribution. In each survey wave, respondents who were currently or recently enrolled in college were also asked to report their major field of study. We use a dummy variable to denote whether the field of study in which the respondent had majored for the longest time by age 29 is a STEM field. College GPA is measured using the respondent's cumulative GPA from the Post-Secondary Transcript Study (PSTRAN). Finally, we include two variables representing the total amounts of loans that the respondent had taken from family and friends and from other sources (including the federal government) to pay for college by age 29. In our analytical sample, some components of the background characteristics (X) and postsecondary characteristics (Z) contain a small fraction of missing values. They are handled by multivariate imputation via chained equations, with ten imputed data sets. The standard errors of our parameter estimates are adjusted for multiple imputation using Rubin's (1987) method.

[Table 2 about here]

Table 2 reports the group-specific means in all variables by gender and race. From the first two panels, we can see that Blacks lag far behind Whites in both educational and labor market outcomes. The degree of Black-White disparity in educational outcomes is similar between men and women (on the probability scale). Compared with White men, Black men are 15 percentage points less likely to have attended a four-year college by age 22, and 17 percentage points less likely to have obtained a BA degree by 29. The corresponding differences are 13 and 18 percentage points for women. The Black-White gap in earnings, by contrast, is much greater among men than among women. In terms of log earnings, the Black-White gap is 0.87 among men but only 0.21 among women. To shed light on the sources of the gender difference, we also report group-specific means in hourly wage and hours worked per year, which suggest that the male Black-White gap in earnings is partly driven by a deficit of Black men in labor force attachment. On average, Black men in our sample worked 1,725 hours per year, nearly 300 hours fewer than did White men. By contrast, Black women in our sample worked 1,600 hours per year, slightly more than White women.

From the third panel of Table 2, we can see large Black-White disparities in socioeconomic background, family structure, and academic achievement. For example, the average parental income among Black respondents is about fifty-two thousand (in 2019 dollars), roughly half that of White respondents, and the average value of parental assets among Blacks is only about 30% as much as that among Whites. Compared with Whites, Black adolescents were also much less likely to live with both biological parents in 1997, much less likely to have a father figure in the household in 1997, and far more likely to have had children by age 18. In the realm of academic achievement, Black respondents in our sample scored substantially lower on the ASVAB test and had poorer high school GPA relative to their White peers. Similar Black-White disparities are evident in contextual characteristics such as college expectations among peers and the school environment.

The last panel of Table 2 shows that even among those who have attended a four-year college, Black students trail their White peers in postsecondary characteristics such as college quality and college GPA. For both men and women, Black college-goers are less likely than their White counterparts to have attended a most competitive, highly competitive, or very competitive college. Moreover, the colleges that Black students attend tend to have lower graduation rates and lower upward mobility rates, i.e., poorer records of lifting low-income students onto the upper rungs of the economic ladder. In addition, Black students on average have a lower college GPA relative to their White peers, which may have contributed to the Black-White gap in BA completion. Among men, Black college-goers are also less likely to have majored in a STEM field. Compared with White students, Black students also tend to have taken more loans from sources other than family and friends.

Results

Observed Black-White Earnings Gaps by Education

Before evaluating the causal effects of college attendance and BA completion, we first describe how the observed Black-White earnings gap varies by education. Table 3 presents average log earnings by gender, race, and education, along with gender-specific Black-White gaps in log earnings both for the full sample and in different educational groups. The first column reproduces the fourth row of Table 2, showing that overall, the Black-White earnings gap is much more pronounced among men than among women. The next four columns show how the average log earnings and the Black-White gap vary across educational groups. We can see that gender differences are substantial, not only in the magnitude of the racial earnings gap but also in the way it varies by education. At each level of education, the racial earnings gap is larger among men than among women. Moreover, the magnitude of the earnings gap differs more sharply between college graduates and the less educated among men than among women. Among men, the earnings gap in log earnings is -0.87 among high school graduates, -0.77 among college dropouts/stopouts, but only -0.2 among college graduates. By contrast, the Black-White gap among women is similar across educational groups — -0.03 among high school graduates, -0.05 among college dropouts/stopouts, and 0.04 among college graduates. None of the within-group earnings gaps for women are statistically distinguishable from zero at the

$p < .05$ level.

[Table 3 about here]

From Table 3, we also find that, in each gender-race group, college graduates earn much more than those with lower levels of education. Among Black men, for example, college graduates have an average of log earnings of 10.67, 0.93 above that of college dropouts/stopouts, and 1.42 above that of high school graduates. The corresponding educational differences among White men are considerably smaller, at 0.36 and 0.75 log points. The steeper gradient associated with a BA degree among Black men leads to a reduction of the racial earnings gap among college graduates, as noted earlier. Such educational gradients, however, should not be interpreted as the causal effects of higher education on earnings. A wide range of pre-college and postsecondary characteristics, such as socioeconomic background, academic preparation, and college GPA, may affect a person's selection into and out of college. Without adjusting for such selection processes, the educational gradients reported in Table 3 are biased estimates of the causal effects of college attendance and BA completion. Moreover, because selection effects may differ by race, the racial differences in these educational gradients do not necessarily reflect racial differences in the economic payoff to college. Below, we turn to our estimates of the causal effects of attending a four-year college, their direct and indirect components, and the implications of these estimates for the Black-White earnings gap.

Total, Direct, and Indirect Effects of College on Earnings

Employing the DML method described previously, we estimate the total effect of attending a four-year college on log earnings, i.e., Δ^g_{tot}, and its direct and indirect components, i.e., Δ^g_{att}, π^g_{comp}, Δ^g_{comp}, Δ^g_{cov}, for each of the gender-race groups. Let us start with the results for men, which are shown in the upper panel of Table 4. From the first column, we can see that the estimated total effect of attending a four-year college is larger for Black men (0.42 log points) than for White men (0.27 log points). This finding is intriguing if we consider that Black men are much less likely than White men to complete college given attendance. From the second and third columns, we can see that the estimated direct and indirect effects of college attendance, i.e., Δ^g_{att} and Δ^g_{ind}, are also greater for Black men than for White men. College attendance per se (short of a BA degree) is expected to boost earnings by 0.22 log points for Blacks and 0.13 log points for Whites. The indirect effect of college, i.e., the effect via a BA degree, is estimated at 0.20 log points for Blacks and 0.14 log points for Whites. None of the Black-White differences in these effects, however, are statistically significant.

[Table 4 about here]

The indirect effect of college (Δ^g_{ind}) is governed by the probability of BA completion given college attendance (π^g_{comp}), the net effect of BA completion on earnings (Δ^g_{comp}), and the covariance between BA completion and its net effect on earnings (Δ^g_{cov}). These components are shown in columns 4-6 of Table 4. On the one hand, we find that given college attendance, Black men are much less likely to complete a BA degree relative to White men. The racial gap in the probability of BA completion is -0.15, constituting a strong stratifying force. On the other hand, the estimated net effect of BA completion on earnings for Black men is 0.79 log points, much higher than that for White men (0.23 log points). The racial difference in the BA completion effect is substantively large and statistically significant, reflecting a strong equalizing effect of a bachelor's degree. Finally, the covariance term also exhibits a racial difference: while it is close to zero for White men, it is estimated at -0.09 for Black men. The latter estimate is substantively significant because the estimated indirect effect of college would have been $0.37 \times 0.79 = 0.29$, instead of 0.20, for Black men if not for the covariance component. As noted earlier, a negative covariance term suggests a pattern of "negative selection" (Brand and Xie 2010): those who would benefit more from a BA degree are less likely to obtain a BA degree given college attendance. Our estimates suggest such a negative selection among Black men but not among White men, adding to the stratifying role of college.

We have seen that among men, the net effect of BA completion is equalizing, i.e., larger among Blacks than among Whites. It is worth asking how this equalizing force would shape the Black-White earnings gap without the stratifying forces associated with racial inequality in degree completion (π^g_{comp}) and in patterns of selection (Δ^g_{cov}). To answer this question, we first assess what we call the *joint effect* of attendance and completion on earnings, i.e., the sum of the direct effect of college attendance and the net effect of BA completion ($\Delta^g_{\text{att}} + \Delta^g_{\text{comp}}$). The results are shown in the last column of Table 4. We find that the joint effect is 1.01 log points among Black men, nearly three times that of White men (0.36 log points). For Black men, the attainment of a BA degree is particularly crucial, accounting for about 80% of the joint effect ($0.79/1.01 = 78.2\%$), compared with 64% for White men ($0.23/0.36 = 63.9\%$).

To see the implications of the racial differences in Δ^g_{att} and Δ^g_{comp} for the Black-White earnings gap, we also estimate a set of potential log earnings for each gender-race group, which reflect average log earnings under hypothetical interventions that set the educational status of a random sample of Blacks and Whites at different levels. The results are shown in Table 5. Echoing our results in Table 4, we find that BA completion has a strong equalizing potential for men. For a random sample of Black and White men, the racial gap in log earnings would be -0.92 if none attended college, -0.78 if everyone attended college (regardless of completion status), but only -0.27 if everyone attended *and* completed college. In other words, the combined effect of college attendance and BA completion on the Black-White earnings gap is about 0.65 log points ($-0.27 - (-0.92) = 0.65$).

[Table 5 about here]

The results for women are shown in the lower panels of Table 4 and Table 5. In contrast to the case of men, we find little evidence of an equalizing effect of BA completion — the estimated effect of BA completion is 0.52 for Black women and 0.55 for White women. Yet, substantial racial inequality exists in the likelihood of BA completion. Given college attendance, the probability of college completion is 0.48 among Black women, compared with 0.68 among White women. On the other hand, both the direct effect of college attendance and the covariance component seem to be equalizing — larger for Black women than for White women (though the differences fall short of reaching conventional levels of statistical significance). As a result of these countervailing forces, the estimated total effect of college attendance on earnings is similar between Black and White women (0.47 versus 0.45 log points).

The lower panel of Table 5 shows the implications of these estimates for the female Black-White earnings gap. For a random sample of Black and White women, the racial gap in log earnings would be -0.14 if none attended college and -0.12 if everyone attended college (regardless of completion status). The gap would be slightly narrower if the stratifying force associated with π^g_{comp} was eliminated — -0.07 if everyone attended college without completing a BA degree, and -0.10 if everyone attended and completed college. Nonetheless, these estimates are all accompanied by large standard errors, and none are statistically significant. Overall, the potential earnings gaps among women are much smaller than those among men, and they do not show much variation by education. Yet, we should note that Black women on average work slightly more than White women (Table 2; see also Chetty et al. 2020b). In fact, the Black-White difference in female employment is even more pronounced among workers with the same level of education (see Table G1). Thus, the relatively small gaps between Black and White women in potential earnings might mask a higher level of racial inequality in potential wages. To explore this possibility, we have conducted parallel analyses for two additional outcomes: hours worked per year and log hourly wage. Results from these analyses are consistent with the above conjecture and are detailed in Supplementary Material G.

In sum, the above results broadly — but not fully — support our hypotheses regarding the dual roles of college summarized in Table 1. While inequality in college completion is present for both men and women ($\pi^{\text{Black}}_{\text{comp}} < \pi^{\text{White}}_{\text{comp}}$), we observe an equalizing role of college only among men. Moreover, it is mostly driven by the net effect of BA completion ($\Delta^{\text{Black}}_{\text{comp}} > \Delta^{\text{White}}_{\text{comp}}$). That said, the equalizing effect of a BA degree for men is substantial: for a random sample of Black and White men, the potential earnings gap would reduce from -0.92 to -0.27 if everyone's educational status was switched from *high school graduate* to *college graduate*. This finding is robust under alternative definitions of college attendance (Supplementary Material C), alternative definitions of BA completion (Supplementary Material D), and alternative measures of earnings (Supplementary Material E). Moreover, as shown in Supplementary Material H, the estimated equalizing effect of a BA degree among men is too large to be plausibly attributed to racial differences in the degree of unobserved selection.

Understanding the Equalizing Effect of a BA Degree

We now conduct a set of additional analyses to better understand the equalizing effect of a BA degree among men. Our earlier discussion suggests that higher education may serve as a "direct equalizer" by helping African American youth circumvent several racialized barriers to economic advancement, such as employer discrimination, neighborhood poverty, and incarceration. On the other hand, given that race is correlated with class background and pre-college academic ability, the equalizing effect of a college degree among men may also reflect a weakened influence of these factors among college graduates. If this is true, a BA degree may be called an "indirect equalizer" as it reduces racial inequality through reducing inequality by class background and pre-college ability. To assess the relative importance of these two mechanisms, we fit and compare three linear models for the (estimated) Neyman-orthogonal signal of each of the potential outcomes (i.e., $Y_i(0,0)$, $Y_i(1,0)$ and $Y_i(1,1)$): a model controlling only for race, a model controlling for race and the percentile rank of parental income, and a model controlling additionally for pre-college academic ability, as measured by the respondent's ASVAB percentile score. By comparing the coefficients of race across these models, we can assess the extent to which the Black-White gaps in potential earnings are explained by class background and pre-college academic ability, and, more importantly, whether the reduced gap at the college graduate level is due to a reduced influence of class and academic backgrounds.

[Table 6 about here]

The results are summarized in Table 6, where the upper, middle, and lower panels correspond to the Black-White gaps in potential log earnings at the levels of *high school graduate*, *college dropout/stopout*, and *college graduate*, respectively. In each panel, we report both the baseline (unadjusted) earnings gap and how it changes after we adjust for parental income and the ASVAB score. The columns titled "explained" show the differences between the unadjusted and adjusted gaps, capturing the explanatory power of parental income and the ASVAB score. Several patterns are noteworthy. First, we find that for men, a BA degree is associated with a much smaller Black-White earnings gap, and this educational difference persists after we control for parental income and the

ASVAB score. In fact, the explanatory power of parental income and the ASVAB score does not vary much by education. At different levels of education, racial differences in parental income translate into a similar amount of the Black-White gap in potential earnings: 0.12 for high school graduates, 0.14 for college dropouts/stopouts, and 0.11 for college graduates. After adjusting for parental income, racial differences in the ASVAB score explain a similar amount of the remaining gap at all levels of education. In combination, parental income and the ASVAB score explain away 0.21 to 0.28 log points of the potential earnings gap between Black and White men. While this amount accounts for the bulk of the earnings gap for men with a BA degree ($0.21/0.25 = 84\%$), it constitutes only about 30% of the earnings gap for less-educated men ($0.25/0.92 = 27\%$ for high school graduates; $0.28/0.83 = 34\%$ for college dropouts/stopouts). Hence, after adjusting for parental income and the ASVAB score, the "residual" earnings gap between Black and White men is 0.67 log points for high school graduates, 0.55 log points for college dropouts/stopouts, but only 0.04 for BA holders.[9] Thus, while racial differences in pre-college class and academic backgrounds are a primary contributor to the Black-White earnings gap among college graduates, they explain only a small fraction of the earnings gaps among less-educated men. This finding suggests that a BA degree is more of a "direct equalizer" than an "indirect equalizer": it narrows the male Black-White earnings gap not by reducing inequality induced by racial differences in class background and pre-college ability, but by lessening the so-called "residual inequality" — a part of inequality more likely driven by labor market factors such as employer discrimination and job access.

Compared with men, the Black-White earnings gap among women is much smaller and does not vary as much by education. For women, the influence of parental income and the ASVAB score appears to be slightly greater at lower levels of education. At a given level of education, racial differences in these pre-college characteristics translate into about 0.32 to 0.46 log points of the potential earnings gap. In contrast to men, this amount can explain away the female Black-White

[9] After parental income rank and the ASVAB score are accounted for, measures of family structure, such as co-residence with both biological parents and presence of a paternal figure, have limited explanatory power. Including them as additional predictors in these models leads to virtually identical results.

gap in potential earnings at all levels of education. In fact, after adjusting for parental income and the ASVAB score, Black women with only a high school diploma are expected to earn significantly more than their White counterparts.

[Table 7 about here]

To put the above findings in perspective, we now turn to parallel results for Hispanic men and women, which are shown in Table 7. We can see that overall, the Hispanic-White earnings gap is smaller than the Black-White earnings gap at almost all levels of education, especially among men. Furthermore, unlike the Black-White earnings gap, the Hispanic-White earnings gap can be largely explained by group differences in parental income and the ASVAB score at all levels of education for both men and women. At the high school level, for example, the estimated residual gap between Hispanic and White men is only 0.02 log points, compared with the 0.67-log-point-gap that separates Black and White men. In sum, our results in Tables 6 and 7 suggest that group differences in pre-college resources and skills are the primary driver of the Hispanic-White earnings gap, the Black-White earnings gap among women, and the Black-White earnings gap among men with a BA degree. Yet, they account for only a small fraction of the massive economic disadvantage faced by less-educated Black men.

Understanding the Black-White Gap in BA Completion

The stratifying role of college is mainly a result of the Black-White disparity in the likelihood of BA completion given attendance. This disparity, as noted earlier, is shaped by a variety of factors, such as racial differences in class background, pre-college academic ability, and college quality. To assess the relative importance of these factors in shaping the Black-White gap in BA completion, we first fit and compare three linear models for the (estimated) Neyman-orthogonal signal of potential BA completion status given college attendance (i.e., $M_i(1)$): a model that controls only for race, a model controlling for race and parental income rank, and a model that controls additionally for the ASVAB score. Since $M_i(0) = 0$ by definition, $M_i(1)$ can be viewed as the causal effect of college attendance on BA attainment for individual i. Thus these models illuminate why the causal effect of college attendance on BA attainment is larger among Whites than among Blacks — specifically, the degree to which it can be attributed to racial differences in class background and pre-college academic ability. The results are shown in the upper panel of Table 8. First, we can see that the bulk of the Black-White gap in BA completion probability among men can be explained by racial differences in parental income. Moreover, after adjusting for the ASVAB score, Black men exhibit a modest, albeit statistically insignificant, advantage over White men in the likelihood of BA completion. Among women, the estimated influence of parental income and the ASVAB score is somewhat weaker. Yet, these two factors still account for about two thirds of the female Black-White gap in BA completion ($0.13/0.20 = 65\%$).

[Table 8 about here]

The above analysis pertains to potential college completion status (given college attendance), i.e., $M_i(1)$., for *all* respondents in our sample. While this population-level analysis aligns with our decomposition of the total college effect (equation 2), it leaves open the question of whether racial differences in college quality contribute independently to the Black-White completion gap among actual college-goers. To address this question, we fit a series of linear regressions only among college-goers, which models the observed BA completion status (M_i) as a function of parental income, the ASVAB score, and all indicators of college quality defined in the Data and Measures section.[10] Results from this analysis, as shown in the lower panel of Table 8, are broadly similar to those based on the full sample. For both men and women, parental income and the ASVAB score can explain most of the Black-White completion gap. Comparing the last two models, we can see that after accounting for racial differences in pre-college class and academic backgrounds, college quality has little independent explanatory power for the Black-White completion gap. This finding echoes previous studies showing that racial gaps in college quality are fully explained or reversed after differences in class and academic backgrounds are taken into account (e.g., Conwell and Quadlin 2021). In sum, the Black-White gap in

[10]The indicators of college quality are computed from administrative data and thus virtually unaffected by the educational and economic outcomes of the NLSY97 respondents.

BA completion is primarily a result of racial disparities in resources and skills formed before college entry.

Policy Implications

We now zoom in on the policy implications of the equalizing and stratifying roles of higher education. As discussed earlier, the equalizing effect of a BA degree is partly offset by the Black-White disparity in college completion, resulting in a modest difference between Black and White men and a similarity between Black and White women in the total effect of college attendance. This finding suggests that a blanket expansion in college enrollment is unlikely to significantly reduce the Black-White earnings gap. As shown in Table 5, for a random sample of Black and White youth, even if everyone attended college, earnings inequality would decline only slightly. On the other hand, given the strong equalizing effect of a BA degree among men, an increase in BA attainment rate may help reduce the male Black-White earnings gap. Thus, it might be supposed that higher education policies aimed at reducing racial inequality should focus on increasing BA completion rates (i.e., graduation rates), especially among Black men. However, BA attainment rate is *the product of college attendance rate and BA completion rate.* Given the current rate of four-year college attendance among Black men (see Table 2), an increase in BA completion rate per se may not substantially change the BA attainment rate in this demographic. From this perspective, both college attendance rate and BA completion rate should be increased to meaningfully boost the proportion of college graduates among Black men. Finally, to the extent that neither college attendance rate nor BA completion rate will reach a point near 100% (at least not in the foreseeable future), part of the Black-White earnings gap will continue to reflect racial disparities in college attendance and completion. Therefore, to reduce racial earnings inequality, higher education policies should also strive to close the Black-White gaps in college attendance and BA completion.

To obtain a more concrete idea of the potential impacts of different policies, we now conduct a thought experiment to predict the counterfactual Black-White earnings gaps under a set of stylized educational interventions. Specifically, we consider three types of hypothetical interventions:

expansion, *redistribution*, and *expansion + redistribution*. By *expansion*, we mean a hypothetical intervention that multiplies everyone's odds of attending/completing college (given their observed characteristics) by a constant such that the overall college attendance/completion rate reaches a prespecified target r. By *redistribution*, we mean a hypothetical intervention that multiplies a person's odds of attending/completing college by a race-specific constant to reach racial parity in college attendance/completion while keeping the overall college attendance/completion rate unchanged. Finally, by *expansion + redistribution*, we mean a hypothetical intervention that multiplies a person's odds of attending/completing college by a race-specific constant such that the college attendance/completion rate reaches a prespecified target r for each racial group. Here, we define these interventions in terms of a proportional increase in everyone's odds of attending/completing college (instead of, for example, an additive/proportional increase on the probability scale) so that it preserves the odds ratio of attending/completing college between individuals, or, in the case of *redistribution* and *expansion + redistribution*, between individuals within the same racial group (Kennedy 2019).

Each of the above interventions can be envisioned for college attendance, BA completion, or both, resulting in nine counterfactuals. For each gender-race group g, we estimate its counterfactual average earnings using the following weighting estimator:

$$\hat{\mathbb{E}}^*[Y|G=g] = \frac{\sum_{G_i=g} w_i Y_i}{\sum_{G_i=g} w_i}, \text{ where } w_i = \eta_i \frac{\overbrace{p^*(A_i|G_i=g,X_i)p^*(M_i|G_i=g,X_i,A_i,Z_i)}^{\text{counterfactual probabilities}}}{\underbrace{\hat{p}(A_i|G_i=g,X_i)\hat{p}(M_i|G_i=g,X_i,A_i,Z_i)}_{\text{factual probabilities}}}. \quad (3)$$

In equation (3), η_i is the NLSY97 sampling weight, $p(\cdot)$ and $p^*(\cdot)$ represent factual and counterfactual probabilities of attending/completing college, and the ratio $p^*(M_i|G_i = g, X_i, A_i, Z_i)/p(M_i|G_i = g, X_i, A_i, Z_i)$ is replaced by 1 if $A_i = 0$ (i.e., if the person did not attend college). This estimator is an extension of Lundberg's (2022) weighting estimator for gap-closing estimands to longitudinal settings. In our context, the counterfactual probabilities $p^*(A_i|G_i = g, X_i)$ and $p^*(M_i|G_i = g, X_i, A_i, Z_i)$ are constructed by transforming the counterfactual odds defined by the intervention back to the probability scale. For *expansion* and *expansion + redistribution*, we set the target attendance/completion rate r at 80% to mimic relatively large-scale interventions.[11] It should be noted that in this analysis, the potential earnings $Y_i(a, m)$ are assumed to be unaffected by these interventions. This assumption will be violated if, for example, an increase in BA attainment rate leads to a more competitive labor market among college graduates, lowering the payoff to a BA degree (Collins 1979; Horowitz 2018). The latter prediction, however, runs counter to empirical trends in the United States, in which the earnings advantage of college graduates has increased despite an expansion in higher education (Bloome et al. 2018), a trend often attributed to the process of skilled-biased technological change (Goldin and Katz 2010). Thus, if history is any guide, higher education expansion will not necessarily lead to a decline in the economic returns to college, which will be shaped by a variety of supply-side, demand-side, and institutional forces. With this caveat in mind, we view the results presented below as a crude but reasonable approximation of the counterfactual earnings gaps that would result in the real economy.

[Table 9 about here]

Results from this counterfactual exercise are shown in Table 9. The first row reproduces the observed earnings gaps shown in Table 3. From the first panel, we can see that a blanket expansion in college attendance would slightly reduce the Black-White earnings gap among men but not among women, as expected. A redistribution in college attendance (without expansion) would reduce the earnings gap by 0.06 log points for both men and women, although this amount constitutes a 30.2% reduction for women but only 7.3% for men. If expansion and redistribution were combined so that the college attendance rate reached 80% for all gender-race groups, the Black-White earnings gap would be reduced by about 13% for both men and women. From the second panel, we can see that

[11] All interventions considered here preserve the conditional distribution of postsecondary characteristics given pre-college characteristics among college-goers, i.e., $p(z|x, A = 1)$. Thus they do not directly interfere with the processes by which college-goers are “matched” to different institutions. By contrast, Chetty et al. (2020a) consider interventions that reallocate college-goers from different income backgrounds into different tiers of colleges but preserve the processes that govern college attendance and BA completion, i.e., interventions that change $p(z|x, A = 1)$ but preserve $p(a|x)$ and $p(m|x, A = 1, z)$. Future research could assess the impacts of more progressive interventions that simultaneously change the processes governing college attendance ($p(a|x)$), college quality ($p(z|x, A = 1)$), and BA completion ($p(m|x, A = 1, z)$).

interventions at the college completion stage would have limited impacts on the Black-White earnings gap, especially for men. This is partly because the current four-year-college attendance rate among Black men is so low that even an increase in BA completion rate to 80% would not substantially alter the educational distribution of this group.

The last panel shows the counterfactual earnings gaps that would result if the three types of interventions were envisioned at both the attendance and completion stages. We can see that an across-the-board increase in both college attendance and BA completion would reduce the male earnings gap by about 20% while leaving the female earnings gap virtually unchanged. These estimates echo our earlier finding that a BA degree is an equalizer among men but not among women. On the other hand, a redistribution in both college attendance rate and BA completion rate (without expansion) would reduce the earnings gap much more among women than among men (in percentage terms). The gender difference in the effects of expansion versus redistribution reveals the different roles of education in shaping the male and female Black-White earnings gaps. Among men, education *moderates* inequality, as a higher level of education, especially a BA degree, leads to a smaller racial earnings gap. Among women, education *mediates* inequality, as a significant part of the overall earnings gap can be removed by eliminating racial inequality in educational attainment. Thus, expansion is more effective at reducing inequality among men but redistribution is more effective at reducing inequality among women. Considering that education also mediates inequality among men, *expansion + redistribution* should be more effective than *expansion* alone at reducing the male earnings gap. This is confirmed in the last row of Table 9: if *expansion + redistribution* was imposed at both stages so that the college attendance rate reached 80% for both Black and White youth with a high school diploma or equivalent, and the BA completion rate reached 80% for both Black and White college-goers, the overall Black-White earnings gap would be reduced by about a third for men and a half for women.

Conclusion

Writing at the climax of the civil rights movement, Otis Dudley Duncan (1968) reasoned that if we could eliminate educational and labor market discrimination against African Americans, the Black-White gap in economic status would “tend to disappear of its own accord” (p. 102). Today, more than half a century past Duncan’s writing, it is disturbingly clear that the gap has shown no signs of disappearance, and, by some indicators, widened (Bayer and Charles 2018). While existing literature has largely focused on forces that maintain and reinforce racial inequality, such as residential segregation and mass incarceration, this study investigates how higher education shapes the Black-White earnings gap. In particular, we highlight the postsecondary system as both an equalizer and a stratifier. Using a novel causal decomposition, a DML estimation method, and data from the NLSY97, we have dissected the total effect of attending a four-year college on earnings into several direct and indirect components. By examining how each of these components differs by race and its correlates, we have isolated the equalizing and stratifying roles of higher education and illuminated their sources.

We find that among men, a BA degree has a strong equalizing effect on earnings, although at the population level, it is partly offset by unequal likelihoods of BA completion and differential patterns of selection. This finding contrasts with recent research on the role of college graduation in the context of intergenerational income mobility, in which the benefit of a BA degree is found to be comparable between students from low- and high-income backgrounds (e.g., Zhou 2019; Fiel 2020; but see Karlson 2019). Thus, our study contributes to the debate on whether a college degree serves as a “great equalizer”: in the context of racial inequality, it still does, albeit only for men.

Why does a college degree reduce the earnings gap between Black and White men but not inequality on other dimensions? Through regression analyses for potential earnings at different levels of education, we find that a BA degree narrows the male Black-White earnings gap primarily by mitigating the “unexplained” penalty of being African American in the labor market, rather than by reducing the influence of class background and pre-college academic ability. This finding reconciles the seemingly inconsistent roles of a BA degree in the contexts of racial inequality versus intergenerational income mobility. It also helps explain why the equalizing effect of a BA degree is restricted to men: for Black and White women, after pre-college class and academic backgrounds are taken into account, there is little residual earnings gap regardless of the level of education, as shown in Table 6.

The above finding prompts the question of *how* a BA degree narrows the "unexplained" inequality for men. Multiple processes may be at work. First, as argued by Arcidiacono et al. (2010), a BA degree allows job seekers to reveal their idiosyncratic abilities in the labor market through information such as grades, majors, and college(s) attended, which may reduce employers' incentives for statistical discrimination. Moreover, given such information, employers should also have less leeway to engage in taste-based discrimination. Relatedly, to the extent that a BA degree is often associated with positive traits such as "hard-working," it may help young Black men counteract many negative stereotypes, such as "unreliable," "scary," or "lacking in work ethic," that they would otherwise suffer (Kirschenman and Neckerman 1991; Moss and Tilly 1996). Second, if young Black men are disproportionately handicapped by neighborhood poverty, which is often associated with limited job opportunities, low return on job referral networks, and a lack of social norms of employment, then a BA degree may narrow the Black-White earnings gap by helping Black men circumvent disadvantaged neighborhoods (Swisher et al. 2013). Finally, given less-educated Black men's disproportionate risk of incarceration and the deleterious effects of incarceration on employment and earnings, a BA degree might also narrow the Black-White earnings gap by reducing racial disparities in incarceration. To be sure, our current analyses do not speak to the relative importance of these processes, and we leave a systematic assessment of them for future research.

To illuminate the policy implications of the equalizing and stratifying roles of higher education, we have considered a series of stylized educational interventions and evaluated the corresponding Black-White earnings gaps under these hypothetical scenarios. Results from this counterfactual analysis suggest that a blanket expansion in college enrollment would not significantly reduce the Black-White earnings gap; nor would an across-the-board increase in college graduation rate per se. If these two expansionary interventions were combined, the Black-White earnings gap could be considerably reduced for men but not for women. To substantially reduce the Black-White earnings gap for both men and women, higher education policies should strive to promote both college attendance and BA completion rates as well as to close racial disparities in these transitions. It should be noted that closing racial disparities in these transitions does not necessarily entail race-conscious interventions (as assumed in our counterfactual analysis) — if we consider that racial disparities in both college attendance and BA completion are largely attributable to racial differences in class background and academic preparation (Ciocca Eller and DiPrete 2018; see also Table 8). Thus, racial disparities in these transitions could also be reduced by race-blind interventions that weaken the influence of class and academic backgrounds on college attendance and degree completion. Such interventions could include personalized outreach efforts that provide counseling and application assistance (e.g., Bettinger et al. 2012; Hoxby et al. 2013), need-based federal, state, and institutional grants (e.g., Alon 2011; Goldrick-Rab et al. 2016), and structured academic and social support during college (e.g., Tinto 2012). Finally, to the extent that job access and employer discrimination play an outsized role in producing racial inequality among men without a college degree, we expect that labor market interventions, such as targeted job creation programs and stricter enforcement of antidiscrimination laws, will be most effective at the lower end of the labor market.

In addition to its substantive contributions and policy implications, this study has employed a new methodological framework for analyzing the effect of higher education on earnings. Unlike the conventional practice of dichotomizing postsecondary attainment into “college-goers” versus “high school graduates” (e.g., Carneiro et al. 2011) or “college graduates” versus “non-graduates” (Brand and Xie 2010), this framework treats BA completion as a mediator that transmits the effect of college attendance on earnings, leading to a causal decomposition that neatly isolates the equalizing and stratifying roles of college. Moreover, to reduce potential model misspecification bias while preserving statistical efficiency, we have used a DML approach to estimate all quantities of interest. Compared with direct applying machine learning algorithms to conventional estimators of causal effects (e.g., propensity score matching), the DML approach provides more robust and efficient estimates along with theoretically valid standard errors.

While our decomposition approach maps more closely than the dichotomous approach onto the sequential process by which people make educational transitions (Mare 1980), it is still an abstraction of the complex and differentiated system of higher education in the US. First, by treating both college attendance and BA completion as binary variables, we have left open the questions of how horizontal stratification by college quality shapes racial earnings inequality, and whether the equalizing and stratifying roles of higher education vary in importance across different types of institutions. The dichotomization of the attendance and completion variables is partly dictated by our data, as the moderate-sized sample of the NLSY97 does not contain enough Black men and women in different types of colleges for a fine-grained analysis. Considering that Black college students tend to attend less selective institutions relative to their White peers and that the value of a BA degree may increase with college selectivity, the equalizing effect of BA completion we found among men may be an underestimate of the equalizing effect of *a BA degree from colleges with similar levels of selectivity*. To test such hypotheses, future research could consider jointly modeling the causal effects of college attendance, college selectivity/quality, and BA completion, and the ways in which they vary by race and its correlates.

Second, this study has focused on the role of four-year institutions, leaving open the question of how the two-year sector of the US postsecondary system shapes the Black-White earnings gap. Future research could adapt our causal diagram and the associated effect decomposition to unpack the economic payoff to attending a two-year college, which comprises not only a direct effect of attendance and an indirect effect via potential attainment of an AA degree, but also an indirect effect via potential transfer to a four-year institution and the associated prospect of attaining a BA degree. Given that two-year colleges currently enroll more than a third of all undergraduate students and that nearly half of all students completing a BA degree had some experience within a two-year institution (Ma and Baum 2016), we consider the relationships between two-year college attendance, educational attainment, and racial economic inequality an important avenue for future research. Finally, this study has also focused on the *gross effect* of BA completion, conflating the direct effect of a BA degree and its “continuation value,” i.e., its effect on earnings via the possibility it creates for attaining even higher levels of education, such as an MA or Ph.D. Given the increasing prevalence of graduate education, more research is needed to investigate how the pursuit and attainment of advanced degrees shape economic inequality (e.g., Torche 2018; Pyne and Grodsky 2020).

Apart from being adapted to study the effects of two-year colleges and other educational transitions, the causal decomposition introduced in this study could also be applied to other domains of inquiry that involve sequential and “state-dependent” mechanisms (Heckman and Borjas 1980; DiPrete and Eirich 2006). For example, in studies of internal labor markets, it could be used to study how early promotions affect career outcomes via the opportunity they create for subsequent promotions to higher levels (e.g., Rosenbaum 1979). In the context of network effects, it could be used to analyze how network access shapes racial inequality in job-search outcomes via racial differences in potential network mobilization given network access (e.g., Pedulla and Pager 2019). Moreover, when studying the socioeconomic consequences of different forms of criminal justice involvement (e.g., Maroto and Sykes 2020), it could be leveraged to isolate the direct effect of conviction from its indirect effect via imprisonment. Given the prevalence of state dependency in social phenomena, we believe that our methodological framework and its variants can find fruitful applications in future research.

References

Adelman, Robert M, Hui-shien Tsao, Stewart E Tolnay, and Kyle D Crowder. 2001. "Neighborhood Disadvantage among Racial and Ethnic Groups: Residential Location in 1970 and 1980." *Sociological Quarterly* 42:603–632.

Alon, Sigal. 2011. "Who Benefits Most from Financial Aid? The Heterogeneous Effect of Need-based Grants on Students' College Persistence." *Social Science Quarterly* 92:807–829.

Alon, Sigal and Marta Tienda. 2005. "Assessing the "Mismatch" Hypothesis: Differences in College Graduation Rates by Institutional Selectivity." *Sociology of Education* 78:294–315.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics* 2:76–104.

Armstrong, Elizabeth A and Laura T Hamilton. 2013. *Paying for the Party*. Cambridge, MA: Harvard University Press.

Athey, Susan and Guido W Imbens. 2019. "Machine Learning Methods that Economists Should Know About." *Annual Review of Economics* 11:685–725.

Attewell, Paul, David Lavin, Thurston Domina, and Tania Levey. 2007. *Passing the Torch: Does Higher Education for the Disadvantaged Pay Off across the Generations?* Russell Sage Foundation.

Autor, David H. 2014. "Skills, Education, and the Rise of Earnings Inequality among the "Other 99 Percent"." *Science* 344:843–851.

Bailey, Martha J and Susan M Dynarski. 2011. "Gains and Gaps: Changing Inequality in US College Entry and Completion." Technical report, National Bureau of Economic Research.

Bayer, Patrick and Kerwin Kofi Charles. 2018. "Divergent Paths: A New Perspective on Earnings Differences between Black and White Men since 1940." *The Quarterly Journal of Economics* 133:1459–1501.

Bertrand, Marianne and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94:991–1013.

Bettinger, Eric P, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *The Quarterly Journal of Economics* 127:1205–1242.

Bjerk, David. 2007. "The Differing Nature of Black-White Wage Inequality across Occupational Sectors." *Journal of Human Resources* 42:398–434.

Black, Dan, Amelia Haviland, Seth Sanders, and Lowell Taylor. 2006. "Why do Minority Men Earn Less? A Study of Wage Differentials among the Highly Educated." *The Review of Economics and Statistics* 88:300–313.

Blau, Peter Michael and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York, NY: Wiley.

Bloome, Deirdre. 2014. "Racial Inequality Trends and the Intergenerational Persistence of Income and Family Structure." *American Sociological Review* 79:1196–1225.

Bloome, Deirdre, Shauna Dyer, and Xiang Zhou. 2018. "Educational Inequality, Educational Expansion, and Intergenerational Income Persistence in the United States." *American Sociological Review* 83:1215–1253.

Boudon, Raymond. 1974. *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York, NY: Wiley.

Bowen, William G, Matthew M Chingos, and Michael S McPherson. 2009. *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton, NJ: Princeton University Press.

Brand, Jennie E and Yu Xie. 2010. "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75:273–302.

Braxton, John M, Amy S Hirschy, and Shederick A McClendon. 1997. "Understanding and Reducing College Student Departure." In *ASHE-ERIC Higher Education Report*, volume 30. San Francisco, CA: Wiley Periodicals.

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101:2754–81.

Charles, Camille Z, Mary J Fischer, Margarita A Mooney, and Douglas S Massey. 2009. *Taming the River: Negotiating the Academic, Financial, and Social Currents in Selective Colleges and Universities*. Princeton University Press.

Cheng, Siwei, Christopher R. Tamborini, ChangHwan Kim, and Arthur Sakamoto. 2019. "Educational Variations in Cohort Trends in the Black-White Earnings Gap among Men: Evidence from Administrative Earnings Data." *Demography* 56:2253–2277.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/Debiased Machine Learning for Treatment and Structural Parameters." *The Econometrics Journal* 21:C1–C68.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020a. "Income Segregation and Intergenerational Mobility across Colleges in the United States." *The Quarterly Journal of Economics* 135:1567–1633.

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2020b. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *The Quarterly Journal of Economics* 135:711–783.

Ciocca Eller, Christina and Thomas A. DiPrete. 2018. "The Paradox of Persistence: Explaining the Black-White Gap in Bachelor's Degree Completion." *American Sociological Review* 83:1171–1214.

Collins, Randall. 1979. *The Credential Society: An Historical Sociology of Education and Stratification*. New York: Academic Press.

Conley, Dalton. 2010. *Being Black, Living in the Red: Race, Wealth, and Social Policy in America*. Berkeley, CA: University of California Press.

Conwell, Jordan A and Natasha Quadlin. 2021. "Race, Gender, Higher Education, and Socioeconomic Attainment: Evidence from Baby Boomers at Midlife." *Social Forces* .

Crowder, Kyle and Scott J South. 2005. "Race, Class, and Changing Patterns of Migration between Poor and Nonpoor Neighborhoods." *American Journal of Sociology* 110:1715–1763.

DiPrete, Thomas A and Gregory M Eirich. 2006. "Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments." *Annual Review of Sociology* pp. 271–297.

Dobbie, Will and Roland G Fryer Jr. 2011. "Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics* 3:158–87.

Duncan, Otis Dudley. 1968. "Inheritance of Poverty or Inheritance of Race?" *On Understanding Poverty* pp. 85–110.

Dwyer, Rachel E, Laura McCloud, and Randy Hodson. 2012. "Debt and Graduation from American Universities." *Social Forces* 90:1133–1155.

Fiel, Jeremy E. 2020. "Great Equalizer or Great Selector? Reconsidering Education as a Moderator of Intergenerational Transmissions." *Sociology of Education* 93:353–371.

Gaddis, S Michael. 2015. "Discrimination in the Credential Society: An Audit Study of Race and College Selectivity in the Labor Market." *Social Forces* 93:1451–1479.

Goldin, Claudia and Lawrence F Katz. 2010. *The Race between Education and Technology*. Harvard University Press.

Goldrick-Rab, Sara, Robert Kelchen, Douglas N Harris, and James Benson. 2016. "Reducing Income Inequality in Educational Attainment: Experimental Evidence on the Impact of Financial Aid on College Completion." *American Journal of Sociology* 121:1762–1817.

Grodsky, Eric. 2007. "Compensatory Sponsorship in Higher Education." *American Journal of Sociology* 112:1662–1712.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning*. Springer Series in Statistics.

Heckman, James J and George J Borjas. 1980. "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence." *Economica* 47:247–283.

Heckman, James J, John Eric Humphries, and Gregory Veramendi. 2018. "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking." *Journal of Political Economy* 126:S197–S246.

Hellerstein, Judith K, David Neumark, and Melissa McInerney. 2008. "Spatial Mismatch or Racial Mismatch?" *Journal of Urban Economics* 64:464–479.

Holzer, Harry J. 1991. "The Spatial Mismatch Hypothesis: What Has the Evidence Shown?" *Urban studies* 28:105–122.

Horowitz, Jonathan. 2018. "Relative Education and the Advantage of a College Degree." *American Sociological Review* 83:771–801.

Hoxby, Caroline, Sarah Turner, et al. 2013. "Expanding College Opportunities for High-achieving, Low Income Students." *Stanford Institute for Economic Policy Research Discussion Paper* 12:014.

Ihlanfeldt, Keith R and David L Sjoquist. 1998. "The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform." *Housing Policy Debate* 9:849–892.

Jack, Anthony Abraham. 2019. *The Privileged Poor: How Elite Colleges are Failing Disadvantaged Students*. Cambridge, MA: Harvard University Press.

Jackson, John W and Tyler J VanderWeele. 2018. "Decomposition Analysis to Identify Intervention Targets for Reducing Disparities." *Epidemiology* 29:825.

Jeffrey, Wesley. 2020. "Crossing the Finish Line? A Review of College Completion Inequality in the United States by Race and Class." *Sociology Compass* 14:e12787.

Jencks, Christopher and Susan E Mayer. 1990. *The Social Consequences of Growing up in a Poor Neighborhood*, pp. 111–86. Washington, DC: National Academy Press.

Jencks, Christopher and Meredith Phillips. 1998. *The Black-White Test Score Gap*. Washington, DC: Brookings Institution Press.

Johnson, William R and Derek A Neal. 1998. "Basic Skills and the Black-White Earnings Gap." In *The Black-White Test Score Gap*, edited by Christopher Jencks and Meredith Philips, pp. 480–97. Washington, DC: Bookings Institution Press.

Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *The quarterly journal of economics* 82:175–197.

Karlson, Kristian Bernt. 2019. "College as Equalizer? Testing the Selectivity Hypothesis." *Social Science Research* 80:216–229.

Kennedy, Edward H. 2019. "Nonparametric Causal Effects Based on Incremental Propensity Score Interventions." *Journal of the American Statistical Association* 114:645–656.

Kirschenman, Joleen and Kathryn M Neckerman. 1991. ""We'd Love to Hire Them, But...": The Meaning of Race for Employers." *The Urban Underclass* 203:203–32.

Kreif, Noémi and Karla DiazOrdaz. 2019. "Machine Learning in Policy Evaluation: New Tools for Causal Inference." In *Oxford Research Encyclopedia of Economics and Finance*.

Lang, Kevin and Michael Manove. 2011. "Education and Labor Market Discrimination." *American Economic Review* 101:1467–96.

Lee, Elizabeth. 2016. *Class and Campus Life: Managing and Experiencing Inequality at an Elite College*. Ithaca, NY: Cornell University Press.

Lochner, Lance and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* 94:155–189.

Lundberg, Ian. 2022. "The Gap-Closing Estimand: A Causal Approach to Study Interventions That Close Disparities Across Social Categories." *Sociological Methods & Research* 0:00491241211055769.

Ma, Jennifer and Sandy Baum. 2016. "Trends in Community Colleges: Enrollment, Prices, Student Debt, and Completion." *College Board Research Brief* 4:1–23.

Manduca, Robert. 2018. "Income Inequality and the Persistence of Racial Economic Disparities." *Sociological Science* 5:182–205.

Mare, Robert D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75:295–305.

Maroto, Michelle and Bryan L Sykes. 2020. "The Varying Effects of Incarceration, Conviction, and Arrest on Wealth Outcomes among Young Adults." *Social Problems* 67:698–718.

Massey, Douglas and Nancy A Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard university press.

Massey, Douglas S, Camille Z Charles, Mary J Fischer, and Garvey Lundy. 2003. *The Source of the River: The Social Origins of Freshmen at America's Selective Colleges and Universities*. Princeton, NJ: Princeton University Press.

Maurin, Eric and Sandra McNally. 2008. "Vive La Revolution! Long-Term Educational Returns of 1968 to the Angry Students." *Journal of Labor Economics* 26:1–33.

McLanahan, Sara and Gary Sandefur. 1994. *Growing up with a Single Parent. What Hurts, What Helps*. Cambridge, MA: Harvard University Press.

Melguizo, Tatiana. 2008. "Quality Matters: Assessing the Impact of Attending More Selective Institutions on College Completion Rates of Minorities." *Research in Higher Education* 49:214–236.

Molina, Mario and Filiz Garip. 2019. "Machine Learning for Sociology." *Annual Review of Sociology* 45:27–45.

Moss, Philip and Chris Tilly. 1996. ""Soft" Skills and Race: An Investigation of Black Men's Employment Problems." *Work and Occupations* 23:252–276.

Mouw, Ted. 2000. "Job Relocation and the Racial Gap in Unemployment in Detroit and Chicago, 1980 to 1990." *American Sociological Review* 65:730–753.

Mouw, Ted. 2002. "Are Black Workers Missing the Connection? The Effect of Spatial Distance and Employee Referrals on Interfirm Racial Segregation." *Demography* 39:507–528.

Moynihan, Daniel Patrick. 1965. *The Negro Family: The Case for National Action*. Washington, DC: US Government Printing Office.

Neal, Derek. 2004. "The Measured Black-White Wage Gap among Women Is Too Small." *Journal of political Economy* 112:S1–S28.

Neal, Derek. 2006. "Why has Black-White Skill Convergence Stopped?" *Handbook of the Economics of Education* 1:511–576.

Neal, Derek A and William R Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104:869–895.

Oliver, Melvin L and Thomas M Shapiro. 2006. *Black Wealth, White Wealth: A New Perspective on Racial Inequality*. New York, NY: Taylor & Francis.

Pager, Devah, Bruce Western, and Bart Bonikowski. 2009. "Discrimination in a Low-Wage Labor Market: A Field Experiment." *American Sociological Review* 74:777–799.

Pattillo, Mary. 2013. *Black Picket Fences: Privilege and Peril among the Black Middle Class*. Chicago, IL: University of Chicago Press.

Pedulla, David S and Devah Pager. 2019. "Race and Networks in the Job Search Process." *American Sociological Review* 84:983–1012.

Pettit, Becky and Bruce Western. 2004. "Mass Imprisonment and the Life Course: Race and Class Inequality in Us Incarceration." *American Sociological Review* 69:151–169.

Polley, EC and M SuperLearner van der Laan. 2017. "Super Learner Prediction: R package, version 2.0-21."

Pyne, Jaymes and Eric Grodsky. 2020. "Inequality and Opportunity in a Perfect Storm of Graduate Student Debt." *Sociology of Education* 93:20–39.

Reardon, Sean F, Rachel Baker, and Daniel Klasik. 2012. "Race, Income, and Enrollment Patterns in Highly Selective Colleges, 1982-2004." .

Reardon, Sean F and Kendra Bischoff. 2011. "Income Inequality and Income Segregation." *American Journal of Sociology* 116:1092–1153.

Reich, Adam D and Seth J Prins. 2020. "The Disciplining Effect of Mass Incarceration on Labor Organization." *American Journal of Sociology* 125:1303–1344.

Robins, James. 1986. "A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period-Application to Control of the Healthy Worker Survivor Effect." *Mathematical Modelling* 7:1393–1512.

Robins, James M. 1997. "Causal Inference from Complex Longitudinal Data." In *Latent Variable Modeling and Applications to Causality*, pp. 69–117. Springer.

Robins, James M, Miguel Angel Hernan, and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11:550–560.

Rosenbaum, James E. 1979. "Tournament Mobility: Career Patterns in a Corporation." *Administrative science quarterly* pp. 220–241.

Royster, Deirdre A. 2003. *Race and the Invisible Hand: How White Networks Exclude Black Men from Blue-Collar Jobs*. Berkeley, CA: University of California Press.

Rubin, Donald B. 1986. "Comment: Which Ifs Have Causal Answers." *Journal of the American statistical association* 81:961–962.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: John Wiley & Sons.

Sakamoto, Arthur, Christopher R Tamborini, and ChangHwan Kim. 2018. "Long-Term Earnings Differentials between African American and White Men by Educational Level." *Population Research and Policy Review* 37:91–116.

Sampson, Robert J. 2012. *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago, IL: University of Chicago Press.

Sampson, Robert J and William Julius Wilson. 1995. *Toward a Theory of Race, Crime, and Urban Inequality*, book section 5, pp. 37–54. Community Justice: An Emerging Field. Lanham, MD: Rowman and Littlefield Publishers, Inc.

Semenova, Vira and Victor Chernozhukov. 2021. "Debiased Machine Learning of Conditional Average Treatment Effects and Other Causal Functions." *The Econometrics Journal* 24:264–289.

Sharkey, Patrick. 2008. "The Intergenerational Transmission of Context." *American Journal of Sociology* 113:931–969.

Sharkey, Patrick. 2013. *Stuck in Place: Urban Neighborhoods and the End of Progress toward Racial Equality*. Chicago, IL: University of Chicago Press.

Small, Mario L and Christopher Winship. 2007. "Black Students' Graduation from Elite Colleges: Institutional Characteristics and Between-institution Differences." *Social Science Research* 36:1257–1275.

Smith, Sandra Susan. 2005. ""Don't Put My Name on It": Social Capital Activation and Job-Finding Assistance among the Black Urban Poor." *American Journal of Sociology* 111:1–57.

Smith, Sandra Susan. 2007. *Lone Pursuit: Distrust and Defensive Individualism among the Black Poor*. New York, NY: Russell Sage Foundation.

Snyder, Thomas D, Cristobal De Brey, and Sally A Dillow. 2019. "Digest of Education Statistics 2017, NCES 2018-070." *National Center for Education Statistics* .

South, Scott J and Kyle D Crowder. 1997. “Escaping Distressed Neighborhoods: Individual, Community, and Metropolitan influences.” *American Journal of Sociology* 102:1040–1084.

Spence, Michael. 1973. “Job Market Signaling.” *The Quarterly Journal of Economics* 87:355–374.

Steele, Claude M. 1988. “The Psychology of Self-affirmation: Sustaining the Integrity of the Self.” *Advances in Experimental Social Psychology* 21:261–302.

Steele, Claude M. 1992. “Race and the Schooling of Black Americans.” *The Atlantic Monthly* 269:68–78.

Steele, Claude M and Joshua Aronson. 1995. “Stereotype Threat and the Intellectual Test Performance of African Americans.” *Journal of Personality and Social Psychology* 69:797–811.

Stevens, Mitchell L, Elizabeth A Armstrong, and Richard Arum. 2008. “Sieve, Incubator, Temple, Hub: Empirical and Theoretical Advances in the Sociology of Higher Education.” *Annu. Rev. Sociol* 34:127–151.

Stuber, Jenny M. 2009. “Class, Culture, and Participation in the Collegiate Extra-curriculum.” *Sociological Forum* 24:877–900.

Stuber, Jenny M. 2011. *Inside the College Gates: How Class and Culture Matter in Higher Education*. Lexington Books.

Swisher, Raymond R, Danielle C Kuhl, and Jorge M Chavez. 2013. “Racial and Ethnic Differences in Neighborhood Attainments in the Transition to Adulthood.” *Social Forces* 91:1399–1428.

Tinto, Vincent. 1994. *Leaving College: Rethinking the Causes and Cures of Student Attrition*. Chicago, IL: University of Chicago Press, 2 edition.

Tinto, Vincent. 2012. *Completing College: Rethinking Institutional Action*. University of Chicago Press.

Tomaskovic-Devey, Donald, Melvin Thomas, and Kecia Johnson. 2005. “Race and the Accumulation of Human Capital across the Career: A Theoretical Model and Fixed-effects Application.” *American Journal of Sociology* 111:58–89.

Torche, Florencia. 2018. “Intergenerational Mobility at the Top of the Educational Distribution.” *Sociology of Education* 91:266–289.

van der Laan, Mark J, Eric C Polley, and Alan E Hubbard. 2007. “Super Learner.” *Statistical Applications in Genetics and Molecular Biology* 6.

Van der Laan, Mark J and Sherri Rose. 2018. *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies*. Springer.

VanderWeele, Tyler J. 2010. “Bias Formulas for Sensitivity Analysis for Direct and Indirect Effects.” *Epidemiology* 21:540.

VanderWeele, Tyler J and Onyebuchi A Arah. 2011. “Bias Formulas for Sensitivity Analysis of Unmeasured Confounding for General Outcomes, Treatments, and Confounders.” *Epidemiology* pp. 42–52.

Voss, Kim, Michael Hout, and Kristin George. 2022. "Persistent Inequalities in College Completion, 1980–2010." *Social Problems* .

Wakefield, Sara and Christopher Uggen. 2010. "Incarceration and Stratification." *Annual Review of sociology* 36:387–406.

Waldinger, Roger. 1993. "The Ethnic Enclave Debate Revisited." *International Journal of Urban and Regional Research* 17:444–444.

Weiss, Andrew. 1995. "Human Capital Vs. Signalling Explanations of Wages." *Journal of Economic Perspectives* 9:133–154.

Western, Bruce. 2002. "The Impact of Incarceration on Wage Mobility and Inequality." *American Sociological Review* 67:526–546.

Western, Bruce. 2006. *Punishment and Inequality in America*. New York, NY: Russell Sage Foundation.

Western, Bruce and Becky Pettit. 2005. "Black-white Wage Inequality, Employment Rates, and Incarceration." *American Journal of Sociology* 111:553–578.

Willis, Robert J. and Sherwin Rosen. 1979. "Education and Self-selection." *Journal of Political Economy* 87:S7–S36.

Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago, IL: University of Chicago Press.

Wilson, William Julius. 1996. *When Work Disappears: The World of the New Urban Poor*. New York, NY: Vintage.

Zhou, Xiang. 2019. "Equalization or Selection? Reassessing the "Meritocratic Power" of a College Degree in Intergenerational Income Mobility." *American Sociological Review* 84:459–485.

Zhou, Xiang. forthcoming. "Attendance, Completion, and Heterogeneous Returns to College: A Causal Mediation Approach." *Sociological Methods & Research* .

Zimmerman, Seth D. 2014. "The Returns to College Admission for Academically Marginal Students." *Journal of Labor Economics* 32:711–754.

Zivich, Paul N and Alexander Breskin. 2021. "Machine Learning for Causal Inference: On the Use of Cross-fit Estimators." *Epidemiology* 32:393–401.

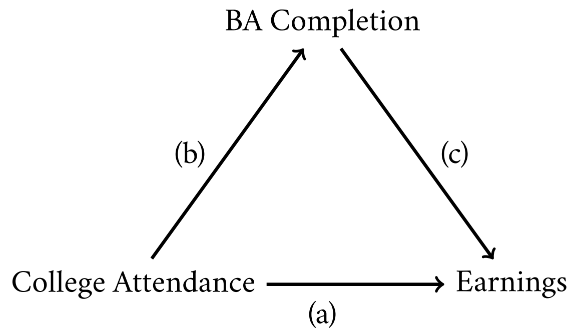

Figure 1: The Effects of College on Earnings in a Causal Diagram
Note: Factors that may confound the relationships between college attendance, BA completion, and earnings are not shown.

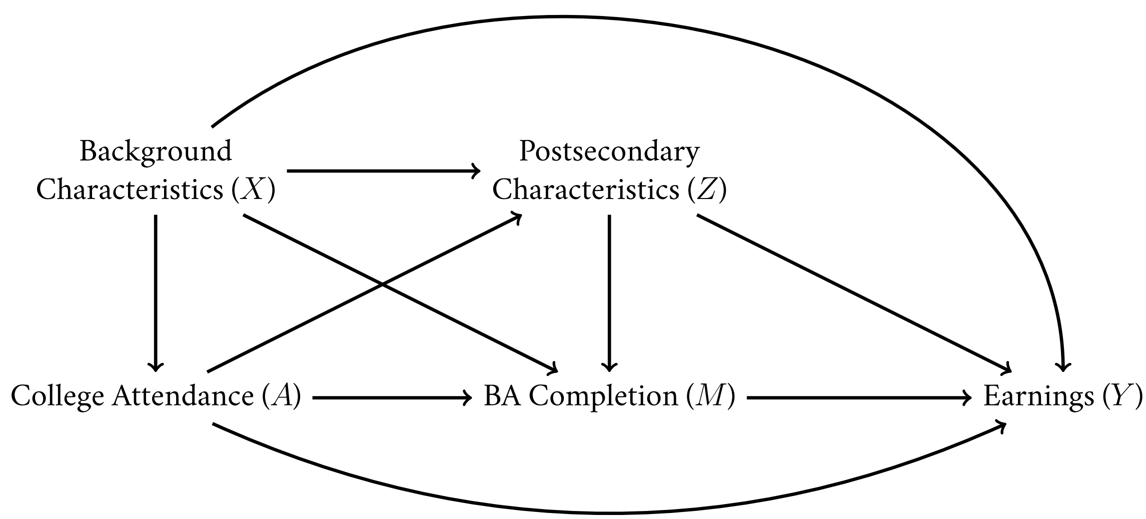

Figure 2: Hypothesized Causal Relationships in a Direct Acyclic Graph

Table 1: Equalizing and Stratifying Roles of Higher Education

	Potential Mechanisms	Empirical Predictions
Equalizing role	Reduction in employer discrimination; heterogeneous effects of college on neighborhood disadvantage and incarceration risk	$\Delta_{\text{att}}^{\text{Black}} > \Delta_{\text{att}}^{\text{White}}$; $\Delta_{\text{comp}}^{\text{Black}} > \Delta_{\text{comp}}^{\text{White}}$
Stratifying role	Racial differences in financial resources, academic preparation, college quality, and psychological processes	$\pi_{\text{comp}}^{\text{Black}} < \pi_{\text{comp}}^{\text{White}}$

Table 2: Group-specific Means in Educational Outcomes, Labor Market Outcomes, Background Characteristics, and Postsecondary Characteristics

		Men		Women	
		Black	White	Black	White
Educational Outcomes	College Attendance by Age 22	0.32	0.47	0.43	0.56
	BA Degree by Age 29	0.16	0.33	0.25	0.43
Labor Market Outcomes	Annual Earnings	31,810	53,922	25,841	34,000
	Log Earnings	9.56	10.43	9.5	9.71
	Hourly Wage	21.97	27.36	18.53	22.83
	Hours Worked per Year	1,725	2,015	1,600	1,535
Background Characteristics	Age at 1997	14.97	14.96	14.96	15
	Parental Years of Schooling	12.78	13.61	12.64	13.54
	Average Parental Income in 1997-2001	52,163	100,027	49,950	96,737
	Parental Assets in 1997	72,379	242,054	66,706	222,994
	Lived with Both Biological Parents	0.29	0.63	0.3	0.59
	Presence of a Father Figure	0.54	0.84	0.56	0.81
	Lived in Rural Area	0.22	0.32	0.17	0.33
	Lived in the South	0.61	0.28	0.63	0.31
	ASVAB Percentile Score	32.03	59.27	35.61	60.67
	High School GPA	2.5	2.88	2.87	3.18
	Substance Use Index	0.84	1.13	0.75	1.13
	Delinquency Index	1.63	1.66	0.91	0.91
	Had Children by Age 18	0.05	0.01	0.16	0.05
	75%+ of Peers Expected College	0.47	0.6	0.5	0.66
	90%+ of Peers Expected College	0.18	0.23	0.22	0.27
	Property Ever Stolen at School	0.31	0.24	0.29	0.19
	Ever Threatened at School	0.22	0.23	0.19	0.18
	Ever in a Fight at School	0.3	0.17	0.15	0.06
Postsecondary Characteristics	Non-profit Private College	0.2	0.22	0.18	0.24
	For-profit College	0.02	0.02	0.05	0.02
	"Most Competitive" College	0	0.03	0.01	0.02
	"Highly Competitive" College	0.04	0.09	0.04	0.06
	"Very Competitive" College	0.07	0.2	0.09	0.22
	Graduation Rate	0.39	0.51	0.41	0.5
	Upward Mobility Rate	0.19	0.27	0.2	0.26
	Majored in STEM	0.23	0.29	0.09	0.1
	College GPA	1.87	2.59	2.13	2.84
	Loans from Family and Friends	514	1,442	616	1,171
	Loans of Other Types	7,513	6,978	9,677	8,241
Sample Size		760	1,614	888	1,588

Note: All statistics are calculated using NLSY97 sampling weights.

Table 3: Black-White Gaps in Observed Log Earnings, Overall and by Level of Education

		Full Sample	HS Graduate	College-goer		
				All	Dropout/Stopout	Graduate
Men	Black	9.56*** (0.06)	9.25*** (0.08)	10.20*** (0.09)	9.74*** (0.12)	10.67*** (0.11)
	White	10.43*** (0.03)	10.12*** (0.05)	10.77*** (0.04)	10.51*** (0.07)	10.87*** (0.05)
	Gap	-0.87*** (0.07)	-0.87*** (0.09)	-0.57*** (0.10)	-0.77*** (0.14)	-0.20 (0.12)
Women	Black	9.50*** (0.05)	9.11*** (0.07)	10.04*** (0.07)	9.54*** (0.11)	10.38*** (0.08)
	White	9.71*** (0.04)	9.14*** (0.06)	10.16*** (0.05)	9.58*** (0.10)	10.34*** (0.05)
	Gap	-0.21** (0.07)	-0.03 (0.09)	-0.12 (0.08)	-0.05 (0.15)	0.04 (0.10)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table 4: Decomposition of the Total Effect of College on Log Earnings by Gender and Race

		Total Effect (Δ^g_{tot})	Direct Effect (Δ^g_{att})	Indirect Effect (Δ^g_{ind})	Completion Prob. (π^g_{comp})	Completion Effect (Δ^g_{comp})	Covariance Term (Δ^g_{cov})	Joint Effect
Men	Black	0.42*	0.22	0.20*	0.37***	0.79***	-0.09*	1.01***
		(0.16)	(0.20)	(0.08)	(0.04)	(0.19)	(0.04)	(0.14)
	White	0.27**	0.13	0.14	0.52***	0.23	0.02	0.36***
		(0.10)	(0.14)	(0.07)	(0.02)	(0.13)	(0.04)	(0.06)
	Diff.	0.14	0.09	0.06	-0.15***	0.56*	-0.12*	0.65***
		(0.19)	(0.24)	(0.11)	(0.04)	(0.23)	(0.06)	(0.15)
Women	Black	0.47***	0.21	0.27***	0.48***	0.52***	0.02	0.73***
		(0.13)	(0.14)	(0.07)	(0.04)	(0.14)	(0.05)	(0.12)
	White	0.45***	0.14	0.32***	0.68***	0.55***	-0.06	0.69***
		(0.10)	(0.10)	(0.06)	(0.02)	(0.10)	(0.03)	(0.09)
	Diff.	0.02	0.07	-0.05	-0.20***	-0.02	0.07	0.04
		(0.16)	(0.17)	(0.10)	(0.04)	(0.18)	(0.06)	(0.16)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table 5: Black-White Gaps in Potential Log Earnings at Different Levels of Education

		HS Graduate	College-goer	College Dropout/Stopout	College Graduate
Men	Black	9.34*** (0.10)	9.76*** (0.14)	9.56*** (0.18)	10.35*** (0.11)
	White	10.27*** (0.04)	10.54*** (0.09)	10.40*** (0.13)	10.63*** (0.05)
	Gap	-0.92*** (0.11)	-0.78*** (0.16)	-0.83*** (0.22)	-0.27* (0.11)
Women	Black	9.30*** (0.08)	9.78*** (0.10)	9.51*** (0.11)	10.03*** (0.10)
	White	9.44*** (0.06)	9.90*** (0.08)	9.58*** (0.09)	10.13*** (0.07)
	Gap	-0.14 (0.10)	-0.12 (0.13)	-0.07 (0.14)	-0.10 (0.12)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table 6: Explaining Black-White Gaps in Potential Log Earnings

	Men		Women	
	Gap	Explained	Gap	Explained
High School Graduate ($Y_i(0,0)$)				
Unadjusted	-0.92*** (0.11)		-0.14 (0.10)	
Adjusted for Parental Income Rank	-0.81*** (0.12)	-0.12*** (0.03)	0.09 (0.10)	-0.23*** (0.04)
Adjusted for Parental Income Rank and ASVAB Score	-0.67*** (0.13)	-0.25*** (0.05)	0.32** (0.11)	-0.46*** (0.05)
College Dropout/Stopout ($Y_i(1,0)$)				
Unadjusted	-0.83*** (0.22)		-0.07 (0.14)	
Adjusted for Parental Income Rank	-0.70** (0.22)	-0.14 (0.07)	0.13 (0.17)	-0.21** (0.07)
Adjusted for Parental Income Rank and ASVAB Score	-0.55* (0.27)	-0.28* (0.12)	0.29 (0.19)	-0.36*** (0.10)
College Graduate ($Y_i(1,1)$)				
Unadjusted	-0.25** (0.09)		-0.10 (0.13)	
Adjusted for Parental Income Rank	-0.14 (0.10)	-0.11*** (0.03)	0.06 (0.13)	-0.16** (0.05)
Adjusted for Parental Income Rank and ASVAB Score	-0.04 (0.11)	-0.21*** (0.05)	0.22 (0.15)	-0.32*** (0.07)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table 7: Explaining Hispanic-White Gaps in Potential Log Earnings

	Men		Women	
	Gap	Explained	Gap	Explained
High School Graduate ($Y_i(0,0)$)				
Unadjusted	-0.18* (0.08)		-0.05 (0.13)	
Adjusted for Parental Income Rank	-0.09 (0.09)	-0.09** (0.03)	0.15 (0.14)	-0.20*** (0.04)
Adjusted for Parental Income Rank and ASVAB Score	-0.02 (0.09)	-0.16*** (0.04)	0.31* (0.15)	-0.37*** (0.05)
College Dropout/Stopout ($Y_i(1,0)$)				
Unadjusted	-0.22 (0.24)		-0.25 (0.22)	
Adjusted for Parental Income Rank	-0.11 (0.25)	-0.10 (0.06)	-0.07 (0.23)	-0.18** (0.06)
Adjusted for Parental Income Rank and ASVAB Score	-0.03 (0.27)	-0.19* (0.08)	0.04 (0.24)	-0.29*** (0.08)
College Graduate ($Y_i(1,1)$)				
Unadjusted	0.01 (0.14)		0.03 (0.20)	
Adjusted for Parental Income Rank	0.09 (0.14)	-0.09*** (0.02)	0.17 (0.20)	-0.14** (0.05)
Adjusted for Parental Income Rank and ASVAB Score	0.15 (0.13)	-0.14*** (0.03)	0.28 (0.20)	-0.26*** (0.06)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table 8: Explaining Black-White Gaps in (Predicted) BA Completion Rates

	Men		Women	
	Gap	Explained	Gap	Explained
Full Sample				
Unadjusted	-0.15*** (0.04)		-0.20*** (0.04)	
Adjusted for Parental Income Rank	-0.03 (0.05)	-0.12*** (0.02)	-0.12* (0.05)	-0.09*** (0.02)
Adjusted for Parental Income Rank and ASVAB Score	0.07 (0.05)	-0.22*** (0.02)	-0.07 (0.05)	-0.13*** (0.02)
College-goers				
Unadjusted	-0.22*** (0.04)		-0.17*** (0.03)	
Adjusted for Parental Income Rank	-0.14*** (0.04)	-0.08*** (0.02)	-0.11*** (0.03)	-0.06*** (0.01)
Adjusted for Parental Income Rank and ASVAB Score	-0.04 (0.04)	-0.18*** (0.02)	-0.06 (0.03)	-0.11*** (0.02)
Adjusted for Parental Income Rank, ASVAB Score, and College Quality	-0.02 (0.04)	-0.19*** (0.02)	-0.05 (0.03)	-0.12*** (0.02)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table 9: Counterfactual Black-White Earnings Gaps under Stylized Educational Interventions

	Men		Women	
	Gap	% Reduced	Gap	% Reduced
Observed Gap	-0.87*** (0.07)		-0.21** (0.07)	
Intervention on College Attendance				
Expansion	-0.79*** (0.09)	9.3%	-0.22** (0.08)	-4.9%
Redistribution	-0.81*** (0.07)	7.3%	-0.15* (0.07)	30.2%
Expansion + Redistribution	-0.75*** (0.10)	13.2%	-0.18* (0.08)	13.4%
Intervention on College Completion				
Expansion	-0.85*** (0.07)	2.3%	-0.20** (0.07)	3.7%
Redistribution	-0.84*** (0.07)	3.5%	-0.17* (0.07)	20.9%
Expansion + Redistribution	-0.84*** (0.08)	2.9%	-0.18** (0.07)	13.8%
Intervention on Both Attendance and Completion				
Expansion	-0.70*** (0.09)	19.6%	-0.20** (0.08)	4.1%
Redistribution	-0.76*** (0.07)	12.2%	-0.10 (0.07)	54.4%
Expansion + Redistribution	-0.58*** (0.10)	33.2%	-0.10 (0.08)	54.2%

Note: *Expansion* means a hypothetical intervention that increases the college attendance/completion rate to 80% in the population. *Redistribution* means a hypothetical intervention that equalizes the college attendance/completion rate between Black and White youth without changing the overall attendance/completion rate in the population. *Expansion + Redistribution* means a hypothetical intervention that increases the college attendance/completion rate to 80% for both Black and White youth. *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are heteroskedasticity-consistent ("sandwich") standard errors.

Supplementary Materials

A Identification and Estimation of Equation (2)

The construction of equation (2) implies that to identify the total effect of college attendance (Δ_{tot}^g) and its various components (Δ_{att}^g, π_{comp}^g, Δ_{comp}^g, Δ_{cov}^g), it suffices to identify the expectations of the following potential outcomes: $\mathbb{E}[Y(0,0)|G=g]$, $\mathbb{E}[Y\big(1,M(1)\big)|G=g]$, $\mathbb{E}[M(1)|G=g]$, $\mathbb{E}[Y(1,0)|G=g]$, and $\mathbb{E}[Y(1,1)|G=g]$, where g denotes a gender-race group (e.g., White women). Here we omit the subscript i for conciseness. Under the sequential ignorability assumption, these quantities are identified via Robins's (1986; 1997) g-formula:

$$\mathbb{E}[Y(0,0)|G=g] = \int \mathbb{E}[Y|x, A=0]dP(x|g), \tag{4}$$

$$\mathbb{E}[Y\big(1,M(1)\big)|G=g] = \int \mathbb{E}[Y|x, A=1]dP(x|g), \tag{5}$$

$$\mathbb{E}[M(1)|G=g] = \int \mathbb{E}[M|x, A=1]dP(x|g), \tag{6}$$

$$\mathbb{E}[Y(1,0)|G=g] = \iint \mathbb{E}[Y|x, A=1, z, M=0]dP(z|x, A=1)dP(x|g), \tag{7}$$

$$\mathbb{E}[Y(1,1)|G=g] = \iint \mathbb{E}[Y|x, A=1, z, M=1]dP(z|x, A=1)dP(x|g), \tag{8}$$

where $P(u|v)$ denotes the cumulative distribution function of U given V. Here the group indicators G are subsumed in X, hence no explicit conditioning on G once X is conditioned on.

To estimate the above quantities, we employ the approach of debiased machine learning (DML; Chernozhukov et al. 2018; Semenova and Chernozhukov 2021), which is characterized by three elements: the construction of a "Neyman-orthogonal signal" for each target parameter (i.e., the components in equation (2) as well as the expected potential outcomes (4)-(8)), a sample-splitting procedure called "cross-fitting," and a linear projection of the Neyman-orthogonal signal for each target parameter on a selected set of "effect modifiers," which, in our case, includes gender, race, and their interaction terms. For each of the expected potential outcomes (4)-(8), we construct a Neyman-orthogonal signal using the (recentered) efficient influence function for its population mean:

$$Y^*(0,0) =: \mathbb{E}[Y|X, A=0] + \frac{1-A}{1-\pi(X)}\big(Y - \mathbb{E}[Y|X, A=0]\big), \tag{9}$$

$$Y^*\big(1, M(1)\big) =: \mathbb{E}[Y|X, A=1] + \frac{A}{\pi(X)}\big(Y - \mathbb{E}[Y|X, A=1]\big), \tag{10}$$

$$M^*(1) =: \mathbb{E}[M|X, A=1] + \frac{A}{\pi(X)}\big(M - \mathbb{E}[M|X, A=1]\big), \tag{11}$$

$$Y^*(1,0) =: \nu_{10}(X) + \frac{A}{\pi(X)}\big(\mu_{10}(X,Z) - \nu_{10}(X)\big) + \frac{A(1-M)}{\pi(X)\big(1-\gamma(X,Z)\big)}\big(Y - \mu_{10}(X,Z)\big), \tag{12}$$

$$Y^*(1,1) =: \nu_{11}(X) + \frac{A}{\pi(X)}\big(\mu_{11}(X,Z) - \nu_{11}(X)\big) + \frac{AM}{\pi(X)\gamma(X,Z)}\big(Y - \mu_{11}(X,Z)\big), \tag{13}$$

where

$$\pi(X) =: \Pr[A=1|X],$$

$$\gamma(X,Z) =: \Pr[M=1|X, A=1, Z],$$

$$\mu_{am}(X,Z) =: \mathbb{E}[Y|X, A=a, Z, M=m],$$

$$\nu_{am}(X) =: \mathbb{E}[\mu_{am}(X,Z)|X, A=a].$$

The Neyman orthogonality of the signals (9)-(13) is given in Chernozhukov et al. (2018) and Van der Laan and Rose (2018). In the above equations, $\mathbb{E}[Y|X, A=a]$, $\mathbb{E}[M|X, A=1]$, $\pi(X)$, $\gamma(X,Z)$, $\mu_{am}(X,Z)$, $\nu_{am}(X)$ are called nuisance functions because they are not of our primary interest but needed for constructing estimators of the target parameters. The signals (9)-(13) are then used to construct the signals for the causal effect parameters Δ^g_{tot}, Δ^g_{att}, π^g_{comp}, and Δ^g_{comp}. For example, the signal for Δ^g_{att} is given by $Y^*(1,0) - Y^*(0,0)$.

The DML approach is then implemented in three steps:

1. Randomly partition the analytical sample $\mathcal{I}$ into J subsamples: $\mathcal{I}_1, \mathcal{I}_2 \dots \mathcal{I}_J$;

2. For units in each subsample $\mathcal{I}_j$ ("estimation sample"), estimate their Neyman-orthogonal

signals using the corresponding nuisance functions estimated from the remainder of the sample ($\mathcal{I} \backslash \mathcal{I}_j$; "training sample");

3. In the full sample, fit a linear regression of each of the estimated signals on gender, race, and their interaction term to estimate group-specific means of the potential outcomes and causal effects. The covariance component Δ^g_{cov} is estimated by $\hat{\Delta}^g_{\text{cov}} = \hat{\Delta}^g_{\text{tot}} - \hat{\Delta}^g_{\text{att}} - \hat{\pi}^g_{\text{comp}} \hat{\Delta}^g_{\text{comp}}$.

In step 2, we estimate each of the nuisance functions using a super learner (van der Laan et al. 2007) composed of Lasso and random forests. Because random forests allow for nonlinear and interaction effects, potential bias due to model misspecification is minimized. Due to the Neyman orthogonality of the signals (9)-(13), our estimates are $\sqrt{n}$-consistent, asymptotically normal, and semiparametric efficient as long as the nuisance function estimates converge to the truth at a faster-than-$n^{-1/4}$ rate, which, unlike the $\sqrt{n}$-consistency required for the nuisance functions in conventional estimators such as inverse probability weighting, is achievable for many machine learning methods. Furthermore, because different subsamples are used for estimating the nuisance functions and for estimating the target parameters, potential bias due to overfitting is also removed. In keeping with Chernozhukov et al. (2018), we use five-fold cross-fitting, meaning that $J = 5$. Standard errors are constructed using the sample variances of the estimated influence functions and adjusted for multiple imputation via Rubin's (1987) method.

B Results from Marginal Structural Models with Inverse Probability Weighting

As noted in the main text and Supplementary Material A, the DML approach to effect estimation is both more robust to model misspecification than parametric methods and more efficient than direct applications of machine learning models to conventional regression- and weighting-based estimators of causal effects. Nonetheless, to illustrate how conventional methods can be used in our setting, we have also estimated the group-specific causal effects Δ^g_{tot}, Δ^g_{att}, π^g_{comp}, Δ^g_{comp} and the potential earnings gaps associated with different levels of education using a more familiar approach: marginal structural models with inverse probability weighting (IPW).

To implement the IPW approach, we first fit two propensity score models, one for college attendance among all respondents and one for BA completion among college-goers. In keeping with conventional practice, we use the logistic regression model to estimate these propensity scores. We then create two inverse-probability-weighted samples: sample 1 and sample 2. In sample 1, the weight for individual i is

$$w_i^1 = w_i^{\text{nlsy97}} \cdot \Big(\frac{A_i \widehat{\Pr}[A_i = 1]}{\hat{\pi}(X_i)} + \frac{(1 - A_i)\widehat{\Pr}[A_i = 0]}{1 - \hat{\pi}(X_i)}\Big),$$

where w_i^{nlsy97} is the NLSY97 sampling weight and $\hat{\pi}(X_i)$ is the estimated propensity score of college attendance. In sample 2, the weight for individual i is

$$w_i^2 = w_i^1 \cdot \Big(1 - A_i + \frac{A_i M_i \widehat{\Pr}[M_i = 1|A_i = 1]}{\hat{\gamma}(X_i, Z_i)} + \frac{A_i(1 - M_i)\widehat{\Pr}[M_i = 0|A_i = 1]}{\hat{\gamma}(X_i, Z_i)}\Big),$$

where $\hat{\gamma}(X_i, Z_i)$ is the estimated propensity score of BA completion given college attendance. The group-specific total effects of college attendance (Δ^g_{tot}) and the corresponding means of potential earnings ($\mathbb{E}[Y(0,0)|G = g]$ and $\mathbb{E}[Y\big(1, M(1)\big)|G = g]$) are then estimated via a weighted linear regression of log earnings on race, college attendance, and their interaction term in sample 1.

Similarly, the group-specific probabilities of BA completion given attendance (π_{tot}^{g}) are estimated via a weighted linear regression of BA completion status on race, college attendance, and their interaction term in sample 1. Finally, the group-specific direct effects of college attendance (Δ_{att}^{g}), net effects of BA completion (Δ_{comp}^{g}), and the corresponding means of potential earnings ($\mathbb{E}[Y(0,0)|G = g]$, $\mathbb{E}[Y(1,0)|G = g]$, $\mathbb{E}[Y(1,1)|G = g]$) are estimated via a weighted linear regression of log earnings on race, college attendance/completion status, and their interaction terms in sample 2. All of these regression models are fit separately for men and for women.

Table B1 shows the IPW estimates of group-specific causal effects (Δ_{tot}^{g}, Δ_{att}^{g}, π_{comp}^{g}, Δ_{comp}^{g}), and Table B2 reports the corresponding estimates of the potential earnings gaps. The IPW estimates of these quantities broadly align with the DML estimates reported in Tables 4 and 5. The estimated patterns of effect heterogeneity are also similar between the two approaches. However, two differences are noteworthy. First, the IPW estimates of the net effect of BA completion (Δ_{comp}^{g}) tend to be greater than the DML estimates. For example, they are 1.12 for Black men and 0.40 for White men, whereas the DML estimates are 0.79 and 0.23, respectively. Second, for all quantities of interest, the IPW estimates are subject to larger standard errors compared with DML. For example, because of the inflated standard errors, the estimated equalizing effect of BA completion among men, i.e., $\hat{\Delta}_{\text{comp}}^{\text{Black men}} - \hat{\Delta}_{\text{comp}}^{\text{White men}}$, is not statistically distinguishable from zero (at the p<.05 level), despite the fact that it is greater in magnitude than the DML estimate (0.72 versus 0.56). These differences suggest that (a) the IPW estimates *may* have suffered from model misspecification bias, and (b) despite our use of parametric models to estimate the propensity scores, IPW is still less efficient compared with DML. These results are consistent with a recent simulation study showing that the DML approach "substantially outperformed all of the other estimators in terms of bias, variance, and confidence interval coverage" in estimating an average causal effect (Zivich and Breskin 2021, p.393).

Table B1: Estimated Causal Effects of College under Marginal Structural Models with Inverse Probability Weighting

		Total Effect	Direct Effect	Completion Prob.	Completion Effect
Men	Black	0.42* (0.21)	0.11 (0.28)	0.36*** (0.05)	1.12*** (0.30)
	White	0.19 (0.22)	0.09 (0.26)	0.48*** (0.04)	0.40 (0.27)
	Diff.	0.23 (0.30)	0.02 (0.38)	-0.12 (0.07)	0.72 (0.40)
Women	Black	0.43** (0.15)	0.15 (0.22)	0.45*** (0.04)	0.54 (0.28)
	White	0.47*** (0.13)	-0.12 (0.20)	0.69*** (0.03)	0.86*** (0.22)
	Diff.	-0.04 (0.20)	0.28 (0.30)	-0.23*** (0.05)	-0.31 (0.35)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are heteroskedasticity-consistent ("sandwich") standard errors.

Table B2: Estimated Potential Earnings Gaps under Marginal Structural Models with Inverse Probability Weighting

		HS Graduate	College-goer	College Dropout/Stopout	College Graduate
Men	Black	9.33*** (0.10)	9.75*** (0.19)	9.44*** (0.26)	10.56*** (0.14)
	White	10.24*** (0.05)	10.43*** (0.21)	10.33*** (0.25)	10.73*** (0.09)
	Gap	-0.91*** (0.11)	-0.68* (0.28)	-0.9* (0.36)	-0.18 (0.17)
Women	Black	9.29*** (0.10)	9.71*** (0.12)	9.44*** (0.20)	9.98*** (0.21)
	White	9.39*** (0.07)	9.86*** (0.12)	9.27*** (0.19)	10.12*** (0.12)
	Gap	-0.11 (0.12)	-0.14 (0.17)	0.17 (0.27)	-0.14 (0.24)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are heteroskedasticity-consistent ("sandwich") standard errors.

C Results under Alternative Definitions of College Attendance

In our main analyses, we use age 22 as the cutoff to define college-goers and non-college-goers (among those who did not hold a BA degree by age 29). By this definition, those who had not attended a four-year college by age 22 or received a BA degree by age 29 are coded as high school graduates, whether they attended a four-year college only after age 22 or ever attended a two-year college. To assess the sensitivity of our results to this measurement choice, we have conducted a series of parallel analyses using alternative age cutoffs for college attendance, where those who attended a four-year college only after the age cutoff ("late college-goers") and those who attended a two-year college but not a four-year college ("two-year college-goers") are either classified as high school graduates ($A = 0$) or excluded from the analyses.

Figure C1 shows our estimates of the Black-White gap in potential log earnings when the age cutoff for college attendance varies from 20 to 25, with late college-goers and two-year college-goers both classified as high school graduates. Figure C2 shows the corresponding estimates when late college-goers are excluded from the analyses, and Figure C3 shows the corresponding estimates when both late college-goers and two-year college-goers are excluded from the analyses. We can see that for both men and women, the potential earnings gaps at different levels of education are similar across alternative age cutoffs, regardless of how late college-goers and two-year college-goers are classified. When both late college-goers and two-year college-goers are excluded from the analyses, the male Black-White gap in potential log earnings exhibits a sharper educational gradient, as it is not only much lower among college graduates than among those with lower levels of education, but also markedly smaller among college dropouts/stopouts than among high school graduates. Our main finding of the equalizing effect of a BA degree among men (and its absence among women) is consistent across all alternative definitions of college attendance.

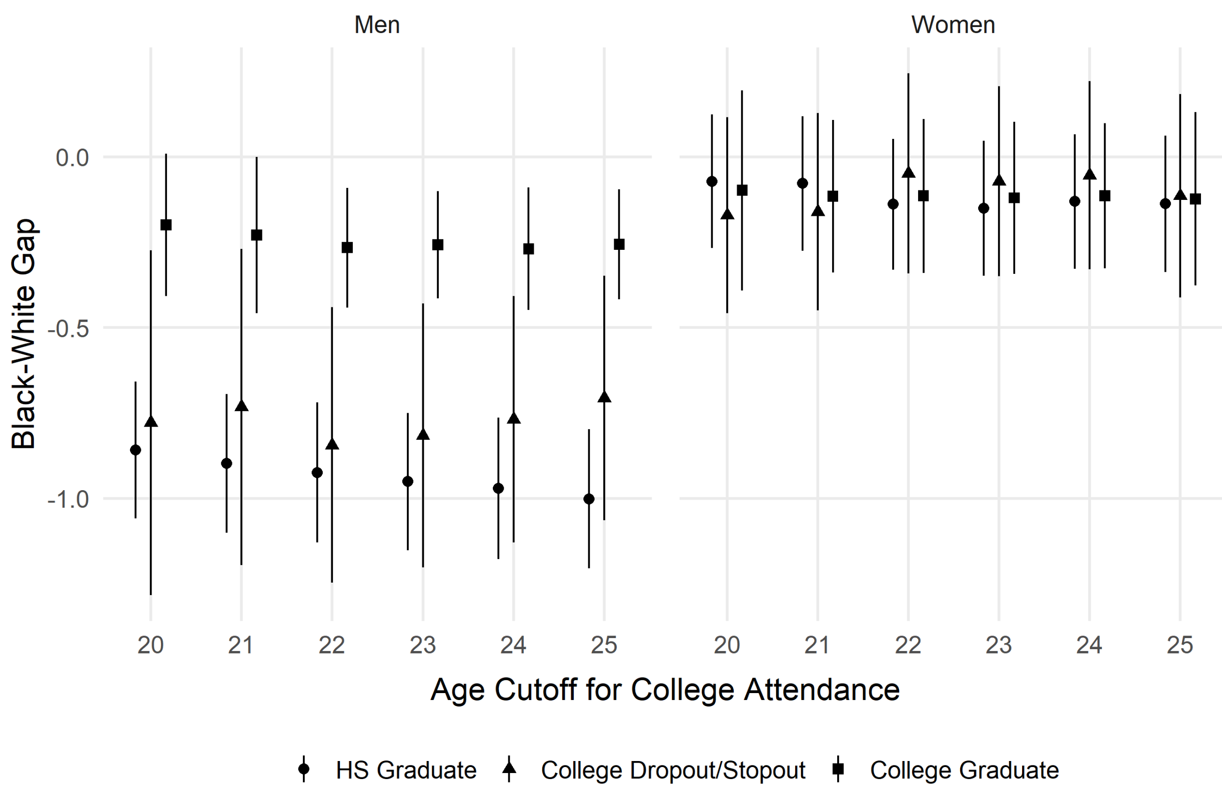

Figure C1: Black-White Gaps in Potential Log Earnings under Different Age Cutoffs for College Attendance

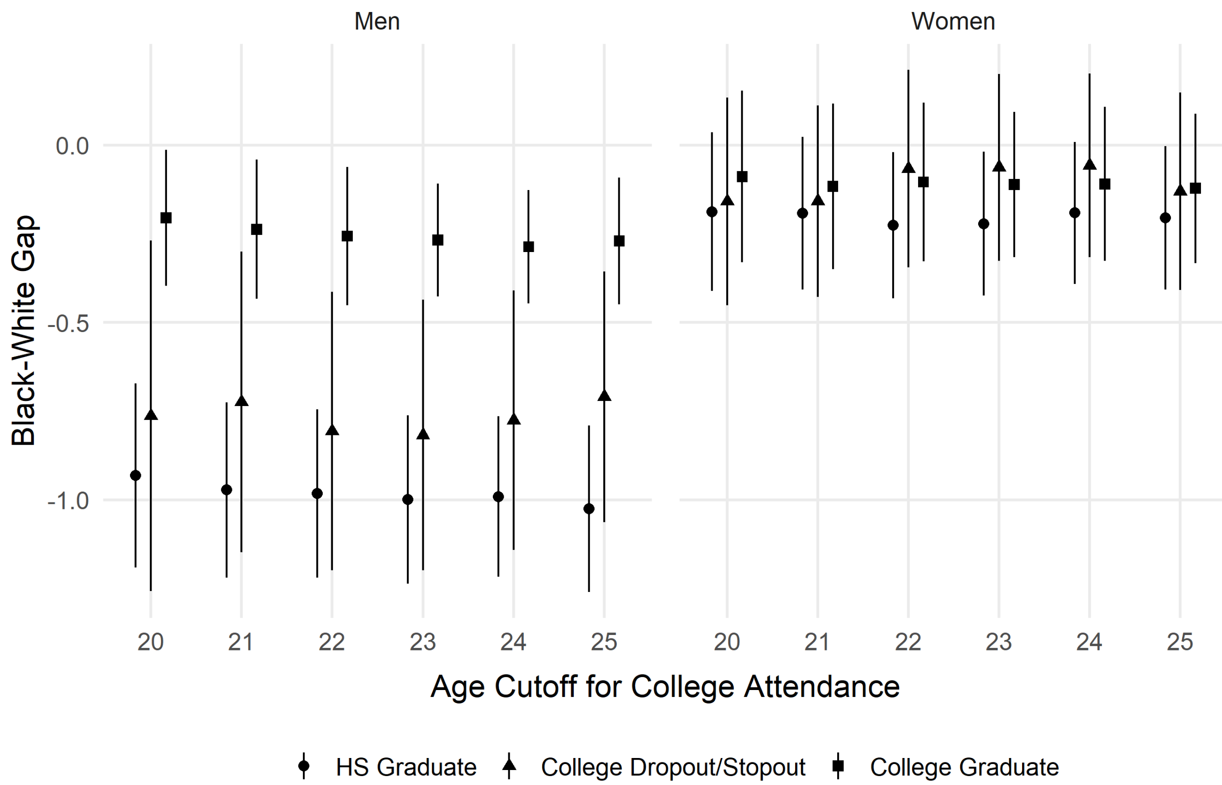

Figure C2: Black-White Gaps in Potential Log Earnings under Different Age Cutoffs for College Attendance with Late College-goers Excluded

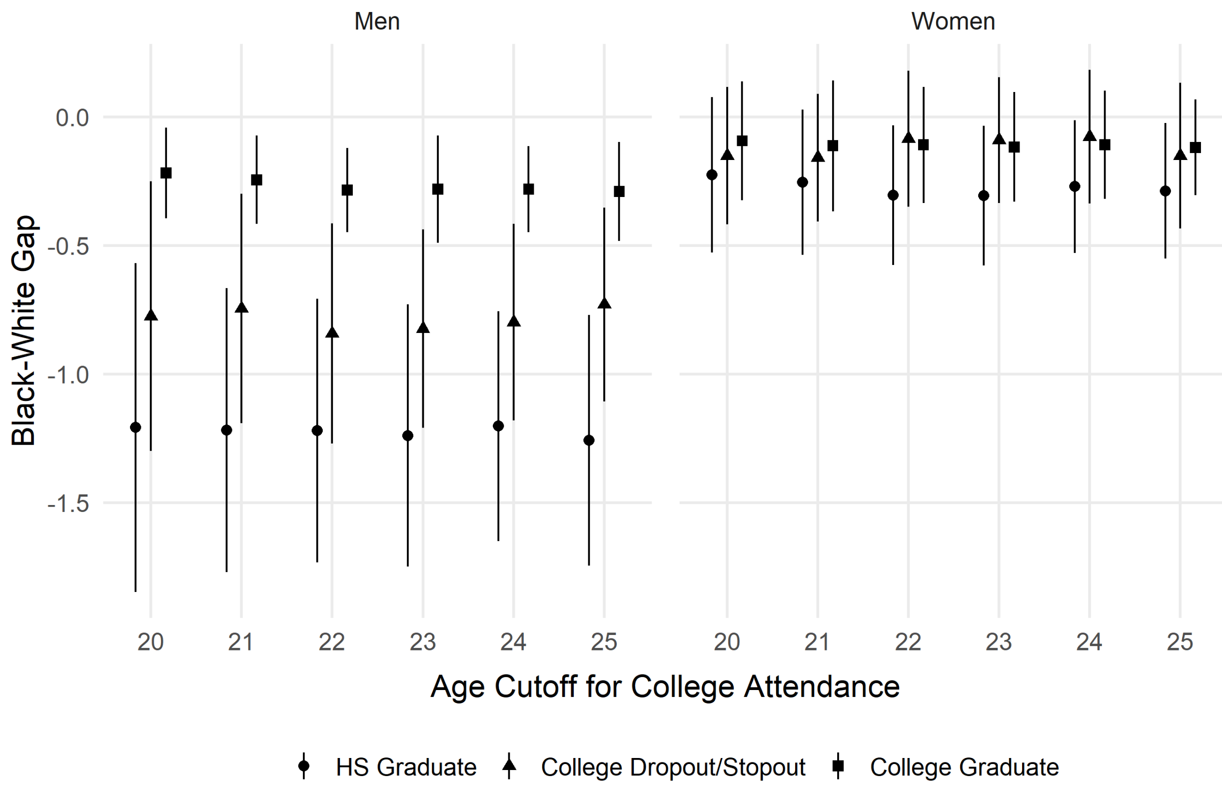

Figure C3: Black-White Gaps in Potential Log Earnings under Different Age Cutoffs for College Attendance with Late College-goers and Two-year College-goers Excluded

D Results under Alternative Definitions of BA Completion

In our main analyses, college graduates are defined as those who had received a BA degree by age 29. We use this age cutoff to accommodate the fact that many young adults obtain their BA degrees beyond the "traditional" ages of college completion, and to minimize misclassification of college graduates as non-graduates when their earnings were measured (at ages 30-33). This definition, however, implies a potentially large variation among college graduates in age at completion. Moreover, Black college graduates tend to complete college at older ages than their White peers. For example, among those who had a BA degree by age 29 in our sample, 30.1% of Black men and 23.8% of Black women completed college at age 26 or older, compared with 17.0% of White men and 15.0% of White women. This difference means that Black college graduates in our sample may have accumulated less work experience than their White peers when their earnings were measured, which should have *inflated*, rather than narrowed, the Black-White earnings gap at the college graduate level. In this regard, our key finding that BA completion narrows the potential earnings gap among men is unlikely a result of racial differences in age at college completion. Nevertheless, to assess the sensitivity of our results to our definition of college graduates, we have conducted two parallel analyses using alternative age cutoffs for BA completion, where those who completed college after the age cutoff ("late college graduates") are either classified as college dropouts/stopouts ($A = 1, M = 0$) or excluded from the analyses.

Figure D1 shows our estimates of the Black-White gap in potential log earnings when the age cutoff for college completion varies from 25 to 29, with late college graduates classified as college dropouts/stopouts. Figure D2 shows the corresponding estimates when late college graduates are excluded from the analyses. We can see that for both men and women, the potential earnings gaps at different levels of education are similar across alternative age cutoffs, regardless of how late college graduates are classified. Our main finding of the equalizing effect of a BA degree among men (and its absence among women) is consistent across all alternative definitions of college completion.

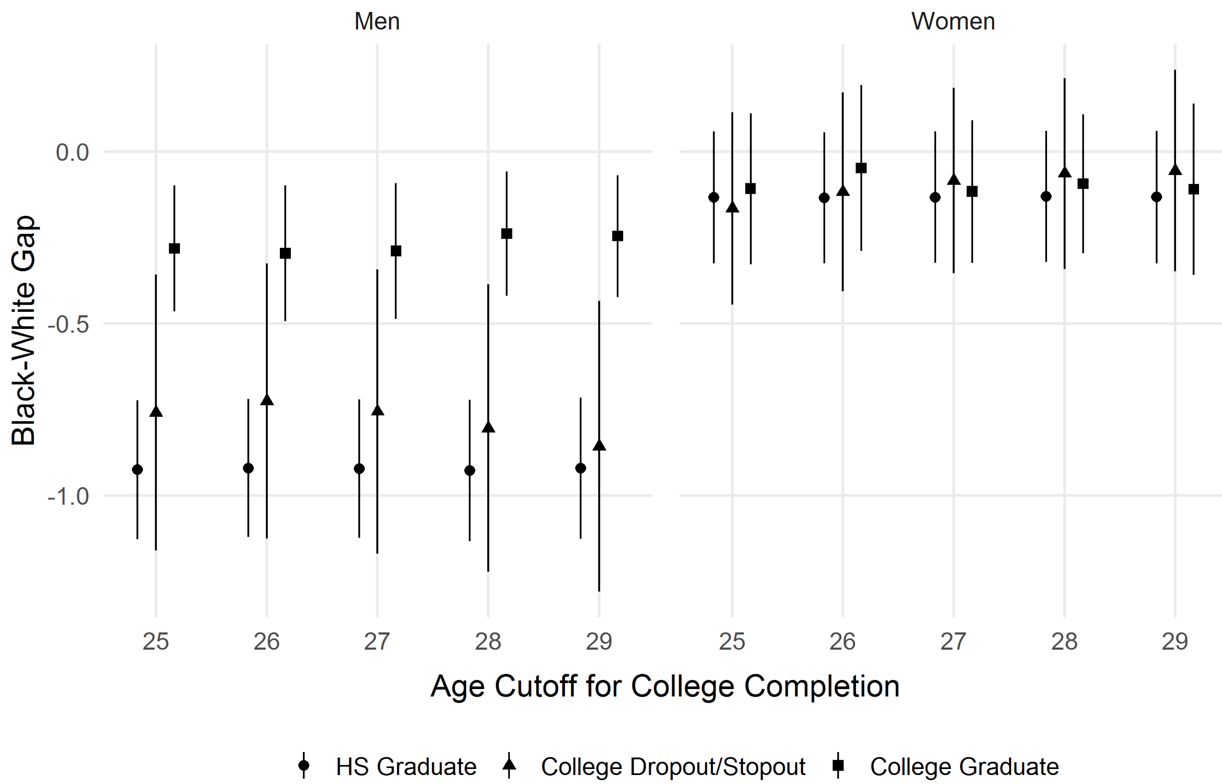

Figure D1: Black-White Gaps in Potential Log Earnings under Different Age Cutoffs for College Completion

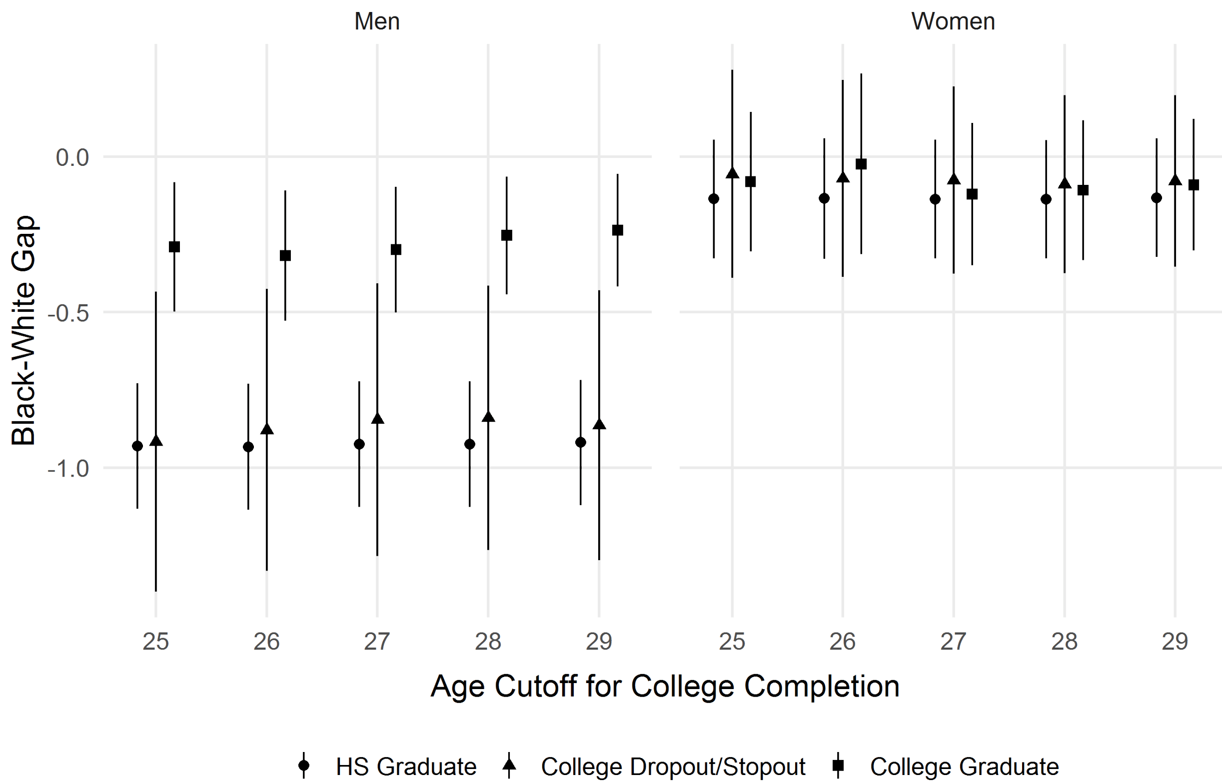

Figure D2: Black-White Gaps in Potential Log Earnings under Different Age Cutoffs for College Completion with Late College Graduates Excluded

E Results under Alternative Outcome Measures

To accommodate individuals with zero earnings, we use $\log(Y + \$1000)$ as our dependent variable in the main analyses. To assess the sensitivity of our results to this measurement choice, we have conducted a series of parallel analyses using alternative adjustments for the log transformation. Figure E1 shows our estimates of the Black-White gap in potential log earnings when the dependent variable is $\log(Y + c)$, where c takes \$1, \$10, \$100, \$1,000, and \$5,000 (in 2019 dollars). We can see that the magnitudes of the earnings gaps in terms of log points are fairly sensitive to the choice of c, especially among less-educated men. This is because less-educated Black men are particularly prone to have zero or very low earnings, making their log earnings relatively sensitive to this adjustment. However, our main finding that the Black-White gap in potential earnings narrows at the BA level among men (but not among women) is unchanged, regardless of the choice of c.

In addition, we have conducted a set of parallel analyses using the percentile rank transformation of earnings (instead of the log transformation), which has recently been used to study racial economic inequality (e.g., Chetty et al. 2020b). The corresponding results are reported in Tables E1-E4, paralleling Tables 3-6 in the main text. We can see that these two sets of results are highly consistent. It should be noted that because a person's earnings rank is a function of both her own earnings and the earnings of everyone else, it depends not only on her own education but also on everyone else's. Thus the stable unit treatment value assumption (SUTVA; Rubin 1986) is mechanically violated. Nonetheless, we can still interpret our estimates of the causal effects of college and the potential earnings gaps from the perspective of local interventions, i.e., interventions applied to a random sample of potential college-goers (Lundberg 2022).

Figure E1: Black-White Gaps in Potential Log Earnings under Alternative Adjustments for Zero Earnings

Table E1: Black-White Gaps in Observed Earnings Ranks, Overall and by Level of Education

		Full Sample	HS Graduate	College-goer		
				All	Dropout/Stopout	Graduate
Men	Black	44.8*** (1.1)	39.2*** (1.3)	56.6*** (1.8)	45.8*** (2.3)	67.9*** (2.4)
	White	62.6*** (0.7)	55.3*** (0.9)	70.8*** (0.9)	63.4*** (1.7)	73.9*** (1.0)
	Gap	-17.8*** (1.3)	-16.1*** (1.6)	-14.2*** (2.1)	-17.6*** (2.9)	-6.1* (2.6)
Women	Black	41.1*** (0.9)	33.2*** (1.0)	51.7*** (1.4)	39.8*** (1.7)	59.9*** (1.7)
	White	47.8*** (0.7)	35.3*** (0.9)	57.5*** (0.9)	44.2*** (1.8)	61.6*** (1.0)
	Gap	-6.7*** (1.1)	-2.1 (1.3)	-5.7*** (1.6)	-4.4 (2.5)	-1.7 (2.0)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table E2: Decomposition of the Total Effect of College on Earnings Rank by Gender and Race

		Total Effect (Δ^g_{tot})	Direct Effect (Δ^g_{att})	Indirect Effect (Δ^g_{ind})	Completion Prob. (π^g_{comp})	Completion Effect (Δ^g_{comp})	Covariance Term (Δ^g_{cov})	Joint Effect
Men	Black	8.2** (2.8)	3.6 (3.2)	4.6** (1.5)	0.37*** (0.04)	17.1*** (3.1)	-1.7* (0.8)	20.6*** (2.5)
	White	7.4*** (2.0)	4.2 (2.4)	3.2* (1.3)	0.52*** (0.02)	5.8** (2.1)	0.2 (0.8)	10.0*** (1.3)
	Diff.	0.8 (3.3)	-0.6 (4.0)	1.5 (1.9)	-0.15*** (0.04)	11.2** (3.7)	-1.9 (1.2)	10.6*** (2.7)
Women	Black	8.3*** (2.0)	3.0 (2.2)	5.4*** (1.1)	0.47*** (0.04)	11.6*** (2.3)	-0.1 (0.7)	14.5*** (2.0)
	White	12.6*** (1.7)	5.0** (1.9)	7.7*** (1.4)	0.68*** (0.02)	12.4*** (1.8)	-0.8 (0.5)	17.4*** (1.6)
	Diff.	-4.3 (2.6)	-2.0 (2.9)	-2.3 (1.8)	-0.21*** (0.04)	-0.8 (2.9)	0.6 (0.9)	-2.8 (2.5)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table E3: Black-White Gaps in Potential Earnings Ranks at Different Levels of Education

		HS Graduate	College-goer	College Dropout/Stopout	College Graduate
Men	Black	40.5*** (1.6)	48.8*** (2.2)	44.1*** (2.7)	61.2*** (1.8)
	White	58.3*** (1.0)	65.8*** (1.7)	62.6*** (2.2)	68.4*** (0.9)
	Gap	-17.8*** (1.9)	-17.0*** (2.8)	-18.5*** (3.5)	-7.2*** (2.0)
Women	Black	37.1*** (1.3)	45.4*** (1.6)	40.1*** (1.8)	51.6*** (1.5)
	White	40.1*** (1.1)	52.8*** (1.3)	45.1*** (1.6)	57.5*** (1.2)
	Gap	-3.0 (1.7)	-7.3*** (2.1)	-5.0* (2.4)	-5.8** (1.9)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table E4: Explaining Black-White Gaps in Potential Earnings Ranks

	Men		Women	
	Gap	Explained	Gap	Explained
High School Graduate ($Y_i(0,0)$)				
Unadjusted	-17.8*** (1.9)		-3.0 (1.7)	
Adjusted for Parental Income Rank	-15.0*** (2.1)	-2.8*** (0.7)	0.3 (1.8)	-3.3*** (0.7)
Adjusted for Parental Income Rank and ASVAB Score	-12.9*** (2.3)	-5.0*** (1.0)	4.1* (1.8)	-7.1*** (0.9)
College Dropout/Stopout ($Y_i(1,0)$)				
Unadjusted	-18.5*** (3.5)		-5.0* (2.4)	
Adjusted for Parental Income Rank	-14.8*** (3.7)	-3.7* (1.5)	-0.9 (2.7)	-4.1*** (1.2)
Adjusted for Parental Income Rank and ASVAB Score	-12.2** (4.5)	-6.2** (2.4)	1.7 (3.0)	-6.7*** (1.6)
College Graduate ($Y_i(1,1)$)				
Unadjusted	-7.2*** (2.0)		-5.8** (1.9)	
Adjusted for Parental Income Rank	-4.5* (2.1)	-2.7*** (0.6)	-3.0 (2.2)	-2.8** (0.9)
Adjusted for Parental Income Rank and ASVAB Score	-2.3 (2.1)	-4.9*** (0.9)	0.6 (2.5)	-6.4*** (1.2)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

F Results from NLSY79 Data

To explore how our findings vary across cohorts and over the life course, we have conducted a supplementary analysis with data from the NLSY79 cohort, using respondents' earnings measured at different ages. In this analysis, college attendance and BA completion are defined in the same way as in our main analyses. However, because of differences between NLSY79 and NLSY97 in survey instruments, the background and postsecondary characteristics we use to adjust for selection are slightly different. Specifically, the background characteristics (X) include gender, race, ethnicity, age at 1979, parental education, parental income, parental occupation, number of siblings, co-residence with both biological parents through age 18, presence of a father figure, rural residence, southern residence, percentile score on the ASVAB test, educational expectation, Rotter's internal–external locus of control scale, whether the respondent had any children by age 18, and three dummy variables indicating whether the respondent's household regularly received magazines, regularly received newspapers, and held a library card when the respondent was at age 14. Because information on college GPA is not available in NLSY79, the postsecondary characteristics (Z) in these analyses include only college type, college quality, field of study, and the total amount of educational loans. College quality is measured using three dummy variables denoting whether the college is one of the "most competitive," "highly competitive," and "very competitive" colleges in Barron's Profile of American Colleges 1986. The measures of college type and field of study are identical to those in our main analysis. To ensure that all background characteristics were measured before college attendance, we restrict our analyses to respondents who were at ages 14-17 in 1979 ($n = 3,659$).

Figure F1 shows our estimates of the Black-White gap in potential log earnings at ages 30-34, 35-39, 40-44, 45-49, and 50-54. Tables F1-F4 report our detailed results when earnings are measured at ages 30-34, paralleling Tables 3-6. We can see that in early adulthood (ages 30-34 and ages 35-39), the estimated male Black-White gap in potential earnings is much lower among college graduates than among those with lower levels of education, echoing our results from the NLSY97. Because few Black men completed college in this sample ($n = 59$), these estimates are accompanied by relatively large standard errors. The equalizing effect of a BA degree seemed to have vanished when the NLSY79 respondents reached their 40s, echoing Tomaskovic-Devey et al.'s (2005) finding that earnings trajectories for members of the NLSY79 cohort in early-to-middle adulthood are flatter for Black and Hispanic men relative to White men, and these disparities are particularly pronounced among the highly educated. However, our estimates suggest a reemergence of the equalizing effect when these respondents reached their 50s (an age group beyond the scope of Tomaskovic-Devey et al.'s analysis), although statistical uncertainty prevents us from drawing a definitive conclusion. In addition, given substantial changes in social and economic conditions, it is difficult to predict whether patterns observed for the earlier cohort will reoccur among members of the NLSY97 cohort. Researchers must follow the NLSY97 respondents as they age to assess whether the equalizing effect of a BA degree persists over their careers.

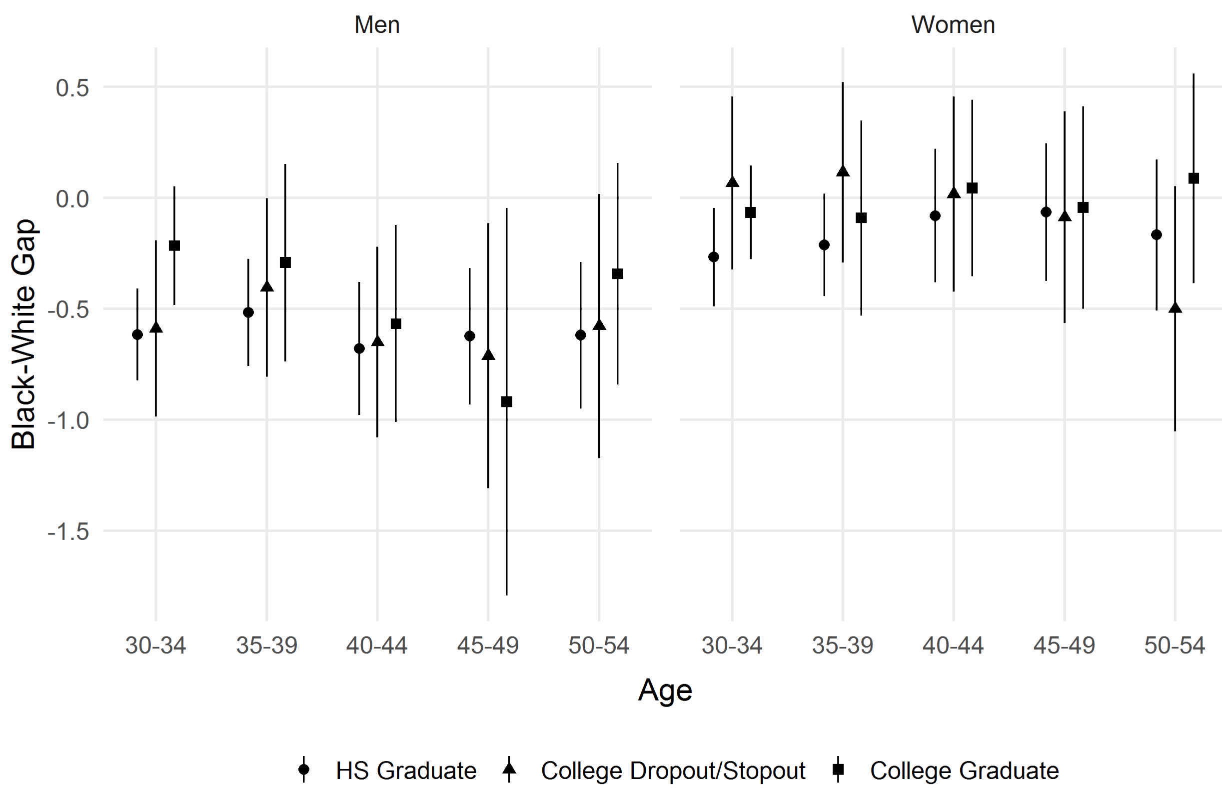

Figure F1: Black-White Gaps in Potential Log Earnings at Different Ages, NLSY79

Table F1: Black-White Gaps in Observed Log Earnings, Overall and by Level of Education, NLSY79 Respondents at Ages 30-34

		Full Sample	HS Graduate	College-goer		
				All	Dropout/Stopout	Graduate
Men	Black	9.98*** (0.06)	9.80*** (0.07)	10.36*** (0.11)	10.02*** (0.15)	10.92*** (0.13)
	White	10.72*** (0.03)	10.48*** (0.05)	11.04*** (0.05)	10.66*** (0.11)	11.18*** (0.05)
	Gap	-0.74*** (0.07)	-0.68*** (0.09)	-0.69*** (0.12)	-0.64*** (0.18)	-0.26 (0.14)
Women	Black	9.61*** (0.06)	9.38*** (0.08)	10.12*** (0.09)	9.88*** (0.13)	10.42*** (0.12)
	White	9.83*** (0.05)	9.64*** (0.06)	10.09*** (0.07)	9.73*** (0.14)	10.23*** (0.09)
	Gap	-0.21** (0.08)	-0.26** (0.09)	0.03 (0.12)	0.15 (0.19)	0.20 (0.15)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table F2: Decomposition of the Total Effect of College on Log Earnings by Gender and Race, NLSY79 Respondents at Ages 30-34

		Total Effect (Δ^g_{tot})	Direct Effect (Δ^g_{att})	Indirect Effect (Δ^g_{ind})	Completion Prob. (π^g_{comp})	Completion Effect (Δ^g_{comp})	Covariance Term (Δ^g_{cov})	Joint Effect
Men	Black	0.42* (0.18)	0.15 (0.21)	0.27* (0.12)	0.29*** (0.07)	0.70** (0.24)	0.07 (0.05)	0.85*** (0.16)
	White	0.33*** (0.08)	0.12 (0.09)	0.21** (0.07)	0.56*** (0.03)	0.33*** (0.10)	0.03 (0.03)	0.45*** (0.09)
	Diff.	0.09 (0.20)	0.03 (0.23)	0.06 (0.13)	-0.27*** (0.08)	0.37 (0.24)	0.04 (0.06)	0.40* (0.17)
Women	Black	0.44** (0.15)	0.26 (0.16)	0.17** (0.06)	0.34*** (0.05)	0.41** (0.15)	0.03 (0.04)	0.67*** (0.13)
	White	0.13 (0.14)	-0.07 (0.17)	0.21** (0.07)	0.55*** (0.03)	0.54*** (0.15)	-0.09 (0.05)	0.47*** (0.10)
	Diff.	0.30 (0.21)	0.33 (0.23)	-0.03 (0.09)	-0.21*** (0.06)	-0.13 (0.21)	0.13* (0.06)	0.20 (0.16)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table F3: Black-White Gaps in Potential Log Earnings, NLSY79 Respondents at Ages 30-34

		HS Graduate	College-goer	College Dropout/Stopout	College Graduate
Men	Black	9.90*** (0.08)	10.32*** (0.16)	10.05*** (0.19)	10.75*** (0.14)
	White	10.52*** (0.06)	10.85*** (0.06)	10.64*** (0.07)	10.97*** (0.06)
	Gap	-0.62*** (0.11)	-0.53** (0.17)	-0.59** (0.20)	-0.22 (0.14)
Women	Black	9.49*** (0.09)	9.93*** (0.12)	9.76*** (0.13)	10.17*** (0.09)
	White	9.76*** (0.07)	9.90*** (0.13)	9.69*** (0.15)	10.24*** (0.07)
	Gap	-0.27* (0.11)	0.03 (0.17)	0.07 (0.20)	-0.07 (0.11)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

Table F4: Explaining Black-White Gaps in Potential Log Earnings, NLSY79 Respondents at Ages 30-34

	Men		Women	
	Gap	Explained	Gap	Explained
High School Graduate ($Y_i(0,0)$)				
Unadjusted	-0.62*** (0.11)		-0.27* (0.11)	
Adjusted for Parental Income Rank	-0.47*** (0.11)	-0.14 (0.08)	-0.07 (0.12)	-0.20* (0.08)
Adjusted for Parental Income Rank and ASVAB Score	-0.42*** (0.12)	-0.20 (0.11)	0.17 (0.13)	-0.44*** (0.10)
College Dropout/Stopout ($Y_i(1,0)$)				
Unadjusted	-0.59** (0.20)		0.07 (0.20)	
Adjusted for Parental Income Rank	-0.53* (0.20)	-0.06 (0.07)	0.33 (0.27)	-0.26 (0.14)
Adjusted for Parental Income Rank and ASVAB Score	-0.36 (0.21)	-0.23* (0.10)	0.64 (0.36)	-0.57** (0.22)
College Graduate ($Y_i(1,1)$)				
Unadjusted	-0.22 (0.14)		-0.07 (0.11)	
Adjusted for Parental Income Rank	-0.06 (0.15)	-0.16* (0.07)	-0.03 (0.13)	-0.04 (0.07)
Adjusted for Parental Income Rank and ASVAB Score	0.08 (0.16)	-0.29** (0.10)	0.02 (0.15)	-0.08 (0.10)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

G Results on Employment and Wages

In our main analyses, we have focused on the Black-White gap in total annual earnings, which consist of wages, salary, as well as income from farms and businesses. Thus, the equalizing effect of a college degree among men may reflect an equalizing effect on working hours, an equalizing effect on wages, an equalizing effect on farm and business income, or a combination thereof. Previous research has shown that the employment rate of less-educated Black men is particularly low, due to limited job access, employment discrimination, and mass incarceration (e.g., Sakamoto et al. 2018). It might then be expected that the equalizing effect of a BA degree on earnings among men is partly, or even primarily, driven by its equalizing effect on employment. On the other hand, we have found that among women, both the observed and the potential earnings gaps are considerably smaller than those among men, and they do not vary much by education. Yet, Black women on average work slightly more than White women; in fact, as shown in Table G1 (last row), the difference between Black and White women's employment is even more pronounced among workers with the same level of education. Thus, the relatively small gaps between Black and White women in potential earnings might mask a greater degree of racial inequality in potential wages. To examine these hypotheses, we have conducted parallel analyses for two additional outcomes: hours worked per year and log hourly wage. For the latter outcome, we restrict our analyses to respondents who engaged in paid work at ages 30-33.

Figure G1 plots our estimates of the direct effects of college attendance, the net effects of BA completion, and the joint effects of attendance and completion on hours worked and log hourly wage. Similar to our analyses on earnings, we also construct a set of potential outcomes and the corresponding Black-White gaps, which are shown in Table G2. Several findings are noteworthy. First, in terms of hours worked, our estimates of the direct effect of college attendance and the net effect of BA completion are both somewhat larger among Black men than among White men. The estimated joint effect of attendance and completion on employment for Black men is 541 hours, compared with 260 hours for White men, which implies an equalizing effect on working hours. As shown in the first panel of Table G2, the estimated male Black-White gap in potential hours worked is over 300 hours among high school graduates, but negligible among college graduates.

Table G1: Black-White Gaps in Observed Hours Worked, Overall and by Level of Education

		Full Sample	HS Graduate	College-goer		
				All	Dropout/Stopout	Graduate
Men	Black	1725*** (41)	1577*** (52)	2036*** (58)	1882*** (87)	2199*** (73)
	White	2015*** (24)	1900*** (36)	2144*** (31)	1987*** (69)	2209*** (32)
	Gap	-290*** (47)	-323*** (63)	-108 (65)	-104 (112)	-10 (79)
Women	Black	1600*** (32)	1370*** (43)	1908*** (40)	1724*** (67)	2035*** (48)
	White	1535*** (23)	1286*** (36)	1728*** (29)	1458*** (60)	1813*** (32)
	Gap	65 (40)	84 (57)	180*** (49)	266** (90)	222*** (58)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

In terms of wages, the estimated direct effect of college attendance is similar by race, but the estimated net effect of BA completion is much stronger among Black men than among White men. As a result, the estimated joint effect of attendance and completion is also greater among Black men than among White men (0.30 log points vs. 0.22 log points). However, the equalizing effect of a college degree reduces but does not eliminate the Black-White wage gap. As shown in the second panel of Table G2, the estimated Black-White gap in potential log hourly wage is -0.22 for high school graduates, -0.29 for college dropouts/stopouts, and -0.14 for college graduates. In sum, our results for men suggest that although both employment and wages contribute to the Black-White earnings gap among those without a BA degree, racial earnings inequality among college graduates is primarily a result of wage inequality.

Among women, the effects of college attendance and BA completion are broadly similar by race for both outcomes. The only exception pertains to the direct effect of college attendance, which appears larger among Black women than among White women (241 hours vs. 11 hours). Because there is little racial difference in hours worked at the high school level, the larger causal effect of college on employment among Black women implies that they would work more than White women

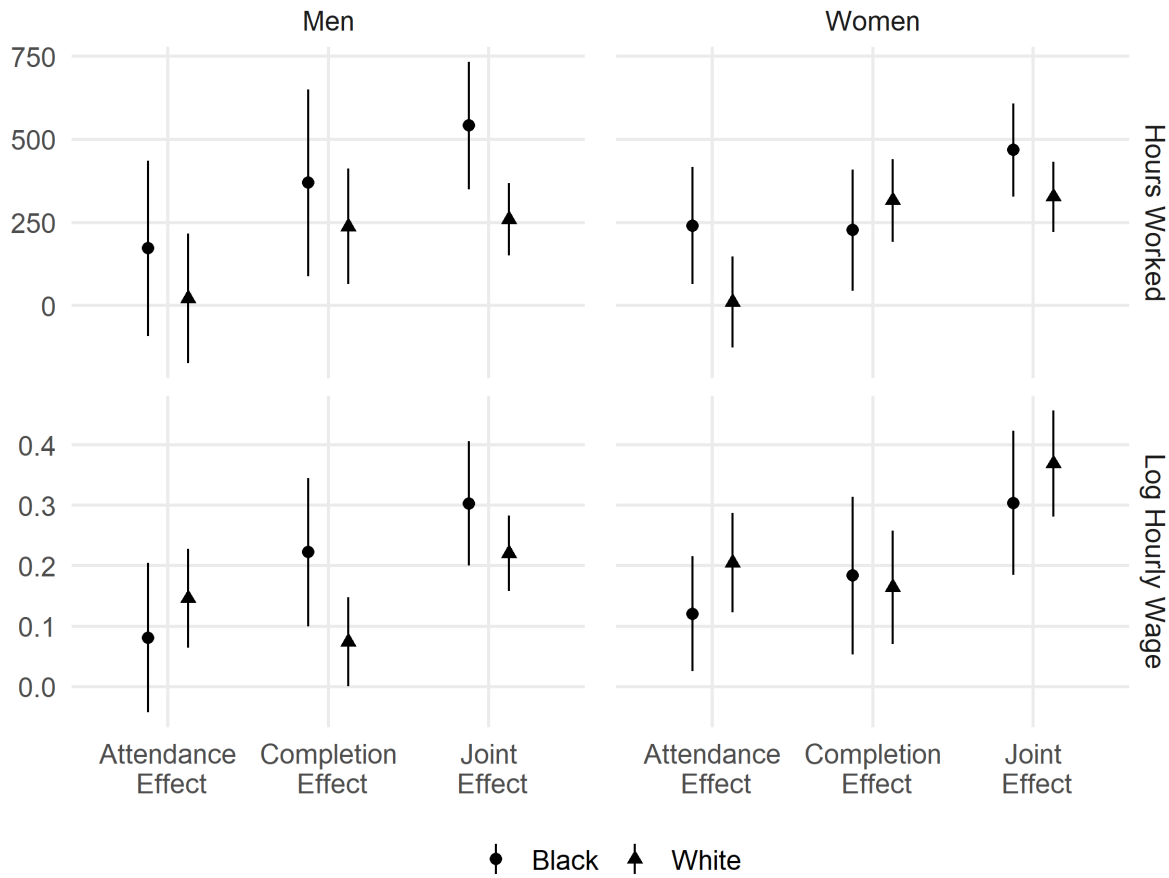

Figure G1: Attendance, Completion, and Joint Effects of College on Hours Worked and Log Hourly Wage

at higher levels of education, as shown in Table G2. In terms of potential wages, however, Black women lag well behind White women at all levels of education. In particular, at the levels of *college dropout/stopout* and *college graduate*, the estimated Black-White gaps in potential log hourly wage are -0.19 and -0.17, larger than the corresponding gaps in potential log earnings (see Table 5). Thus, among women, the relatively mild earnings gaps reported earlier mask sizable gaps in wages, which are only partly compensated by the slightly longer hours worked by Black women. Unlike the case of men, neither college attendance nor BA completion seems to have an equalizing effect on women's wages.

Table G2: Black-White Gaps in Potential Hours Worked and Log Hourly Wage

			HS Graduate	College Dropout/Stopout	College Graduate
Men	Hours Worked	Black	1584*** (59)	1757*** (122)	2126*** (76)
		White	1888*** (39)	1910*** (90)	2148*** (39)
		Gap	-304*** (71)	-154 (156)	-22 (76)
	Log Wage	Black	2.73*** (0.03)	2.82*** (0.05)	3.04*** (0.04)
		White	2.96*** (0.02)	3.10*** (0.03)	3.18*** (0.02)
		Gap	-0.22*** (0.04)	-0.29*** (0.06)	-0.14** (0.04)
Women	Hours Worked	Black	1452*** (45)	1692*** (77)	1919*** (55)
		White	1407*** (39)	1418*** (56)	1735*** (38)
		Gap	44 (59)	274** (97)	185** (68)
	Log Wage	Black	2.63*** (0.03)	2.75*** (0.04)	2.93*** (0.06)
		White	2.73*** (0.03)	2.94*** (0.03)	3.10*** (0.04)
		Gap	-0.10** (0.04)	-0.19*** (0.05)	-0.17** (0.06)

Note: *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors.

H Sensitivity Analyses

H.1 Rationale and Results

Our identification of the total, direct, and indirect effects of college on earnings rests on the strong and untestable assumption that no unobserved confounders exist for the relationships between college attendance, BA completion, and earnings. We conduct a sensitivity analysis to explore the degree to which our estimates of the total effect of college (Δ^g_{tot}), the direct effect of attendance (Δ^g_{att}), and the net effect of completion (Δ^g_{comp}) are robust to unobserved confounding. Specifically, we employ the bias factor approach developed by VanderWeele (2010) and VanderWeele and Arah (2011). Because our estimates of the college effects are comparable by gender and they differ substantially by race only among men, we focus on men in this analysis.

First, although we have adjusted for an array of background characteristics in our analyses, there may still be unobserved individual attributes that affect both college attendance (A) and log earnings (Y). For analytical tractability, we consider a binary unobserved confounder U, say a strong interpersonal skill, that affects both college attendance and earnings. Under some simplifying assumptions on the homogeneity of the relationships between U, A, and Y (see Section H.2), the bias for the estimated Δ^g_{tot} is given by

$$\text{bias}[\Delta^g_{\text{tot}}] = \alpha^g_{\text{tot}}\beta^g_{\text{tot}}, \tag{14}$$

where α^g_{tot} denotes the difference in the prevalence of U between high school graduates ($A = 0$) and college-goers ($A = 1$) given baseline covariates X, and β^g_{tot} is the average difference in log earnings between those with and without U given college attendance status A and baseline covariates X.

Given the multiplicative structure of the bias formula (14), let us use $\gamma^g_{\text{tot}} = \text{sign}(\alpha^g_{\text{tot}}\beta^g_{\text{tot}})\sqrt{\alpha^g_{\text{tot}}\beta^g_{\text{tot}}}$, i.e., the geometric mean of α^g_{tot} and β^g_{tot}, as a measure of the strength of unobserved selection in group g. Figure H1 shows the bias-adjusted estimates of Δ^g_{tot} for Black and White men across a range of possible values of γ^g_{tot}. We set the maximum of γ^g_{tot} at 0.3, corresponding to an extreme level of unobserved selection, which would arise, for example, when the unobserved confounder U increases earnings by 0.3 log points conditional on X and A and its prevalence differs by 30 percentage points between high school graduates and college-goers conditional on X. Given that an unobserved characteristic that boosts earnings is likely also positively associated with college attendance (and vice versa), we may focus on the right part of Figure H1, where $\gamma^g_{\text{tot}} > 0$. We can see that in this case, although our estimates of Δ^g_{tot} will be upwardly biased, they are quite robust to unobserved confounding for both racial groups. For example, if the unobserved characteristic increases earnings by 0.2 log points given X and A and its prevalence differs by as much as 25 percentage points between high school graduates and college-goers given X, then $\gamma^g_{\text{tot}} = \sqrt{0.2 \times 0.25} = 0.22$. In this case, the bias-adjusted estimate of the total effect stands at 0.37 log points among Black men and 0.22 log points among White men, which are fairly close to the baseline estimates (0.42 and 0.27). Moreover, as shown by the confidence bands, almost all of the bias-adjusted estimates in Figure G1 are statistically significant at the $p < .05$ level.

Second, unobserved confounders may exist for the causal effect of BA completion (M) and log earnings (Y). In this case, while the total effect of college attendance may still be unbiased, the direct effect of college attendance (Δ^g_{att}) and the net effect of BA completion (Δ^g_{comp}) will likely be over- or under-estimated. To explore the direction and magnitude of potential bias, we again consider a binary unobserved confounder U, say a positive peer influence, that affects both BA completion and earnings and may itself be affected by college attendance A. Then, under some simplifying assumptions on the homogeneity of the relationships between U, A, M, Z and Y (see Section H.2), the biases for estimated Δ^g_{att} and Δ^g_{comp} are given by

$$\text{bias}[\Delta^g_{\text{att}}] = -\pi^g_{\text{comp}}\alpha^g_{\text{comp}}\beta^g_{\text{net}}, \tag{15}$$

$$\text{bias}[\Delta^g_{\text{comp}}] = \alpha^g_{\text{comp}}\beta^g_{\text{net}}, \tag{16}$$

where π^g_{comp} is the probability of BA completion given college attendance among members of group g (see equation 2), α^g_{comp} denotes the difference in the prevalence of U between college

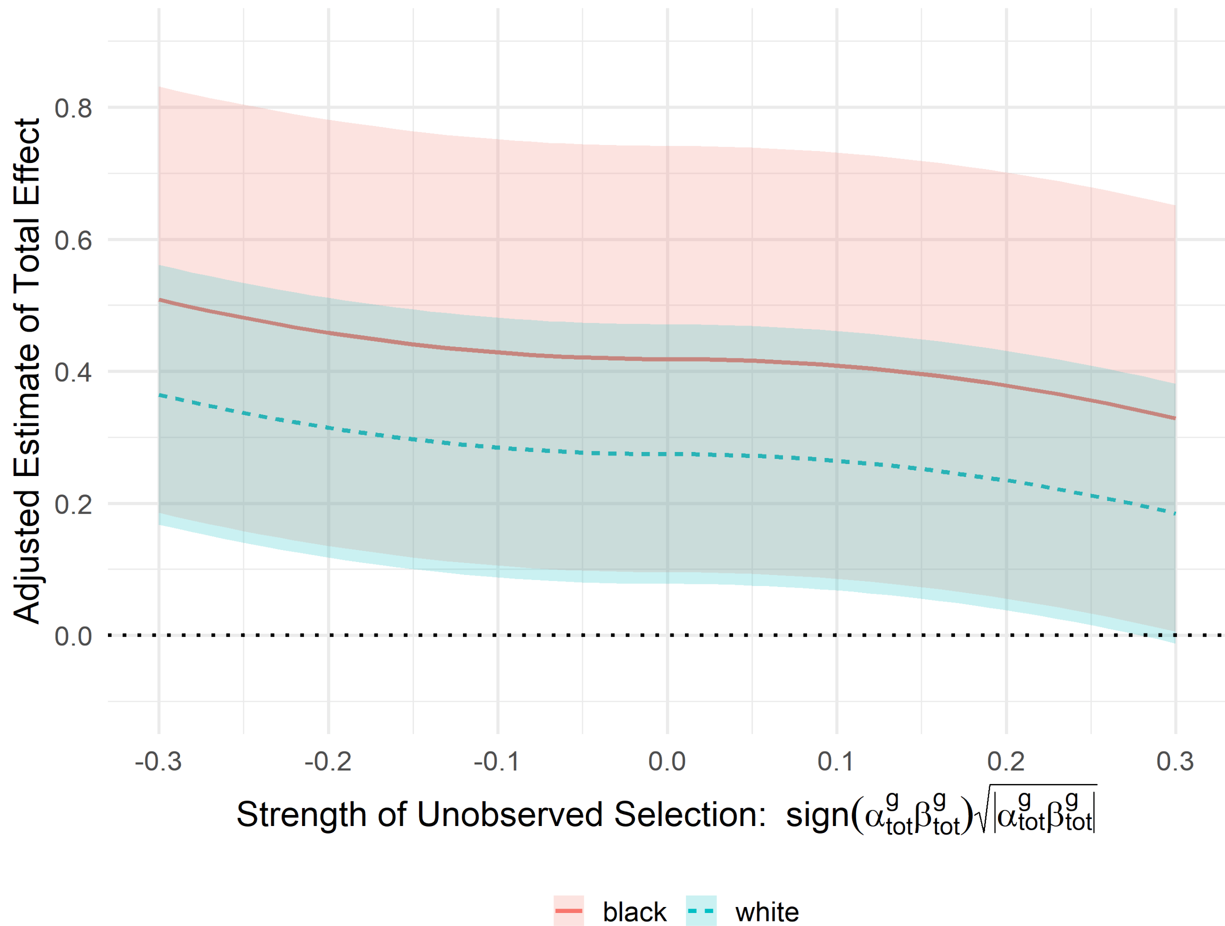

Figure H1: Sensitivity Results for the Total Effect of College (Δ^g_{tot}) on Men's Earnings

dropouts/stopouts ($A = 1, M = 0$) and college graduates ($A = M = 1$) given both baseline and postsecondary characteristics (X and Z), and β^g_{net} denotes the net difference in log earnings between those with and without the unobserved characteristic U given X, A, M, and Z.

As before, given the multiplicative structure of the bias formulas (15) and (16), we use $\gamma^g_{\text{net}} = \text{sign}(\alpha^g_{\text{comp}}\beta^g_{\text{net}})\sqrt{\alpha^g_{\text{comp}}\beta^g_{\text{net}}}$, i.e., the geometric mean of α^g_{comp} and β^g_{net}, as a measure of the strength of unobserved selection in group g. Figure H2 reports the bias-adjusted estimates of Δ^g_{att} and Δ^g_{comp} across a range of possible values of γ^g_{net}. When assessing the bias-adjusted estimates of Δ^g_{att}, we substitute our DML estimate $\hat{\pi}^g_{\text{comp}}$ for π^g_{comp}. The standard errors of the bias-adjusted estimates of Δ^g_{att} are adjusted accordingly.[12] Since an unobserved characteristic that boosts earnings is likely

[12]Specifically, since the bias-adjusted estimate of Δ^g_{att} is $\hat{\Delta}^g_{\text{att}} + \hat{\pi}^g_{\text{comp}}(\gamma^g_{\text{net}})^2$, its standard error is

positively associated with BA completion, it is reasonable to assume that α^g_{comp} is also positive. Thus, we may focus on cases where $\gamma^g_{\text{net}} > 0$. In this case, the direct effect of attendance (Δ^g_{att}) is likely underestimated and the net effect of BA completion (Δ^g_{comp}) is likely overestimated. For Δ^g_{att}, we can see that our point estimates are not much affected by unobserved selection, although they are not statistically significant regardless of the level of bias (see also Table 4). For Δ^g_{comp}, our estimate for White men is also not statistically distinguishable from zero at most levels of γ^g_{net}. However, our estimate of Δ^g_{comp} for Black men is highly robust. For example, even if the unobserved characteristic increases earnings by 0.3 log points (given X, A, M, and Z) and its prevalence differs by as much as 30 percentage points between college dropouts/stopouts and college graduates (given X and Z) — in which case $\gamma^g_{\text{net}} = 0.3$, the bias-adjusted estimate of Δ^g_{comp} for Black men still stands at 0.73 log points.

Moreover, it is clear that if the sensitivity parameters are identical between Blacks and Whites, the same amount of bias will afflict our estimates of Δ^g_{tot} and Δ^g_{comp} for both groups, leaving our finding about racial differences in college effects unchanged. However, the sensitivity parameters may differ by race. For example, peer influence might be more crucial in shaping college completion and earnings among Blacks than among Whites. In this case, the larger estimate of Δ^g_{comp} among Black men could be a result of differential selection bias, which can be gauged by $\alpha^{\text{Black men}}_{\text{comp}}\beta^{\text{Black men}}_{\text{net}} - \alpha^{\text{White men}}_{\text{comp}}\beta^{\text{White men}}_{\text{net}}$. Nonetheless, given our estimates of the BA completion effect for Black and White men (0.79 versus 0.23), the differential selection bias would have to reach 0.56 to explain away the racial difference. Considering the range of plausible values for our sensitivity parameters and the associated bias, it is highly unlikely that unobserved confounding alone can account for the equalizing effect of a college degree we find among men. The robustness of the equalizing effect can be seen from the fact that the solid line (for Black men) is *uniformly* above the dashed line (for White men) in the right panel of Figure H2.

$\sqrt{\text{var}[\hat{\Delta}^g_{\text{att}}] + (\gamma^g_{\text{net}})^4\text{var}[\hat{\pi}^g_{\text{comp}}] + 2(\gamma^g_{\text{net}})^2\text{cov}[\hat{\Delta}^g_{\text{att}}, \hat{\pi}^g_{\text{comp}}]}$. For the latter quantity, we use a plug-in estimator where $\text{var}[\hat{\Delta}^g_{\text{att}}]$, $\text{var}[\hat{\pi}^g_{\text{comp}}]$, and $\text{cov}[\hat{\Delta}^g_{\text{att}}, \hat{\pi}^g_{\text{comp}}]$ are estimated via the sample variances and covariances of the estimated influence functions of $\hat{\Delta}^g_{\text{att}}$ and $\hat{\pi}^g_{\text{comp}}$.

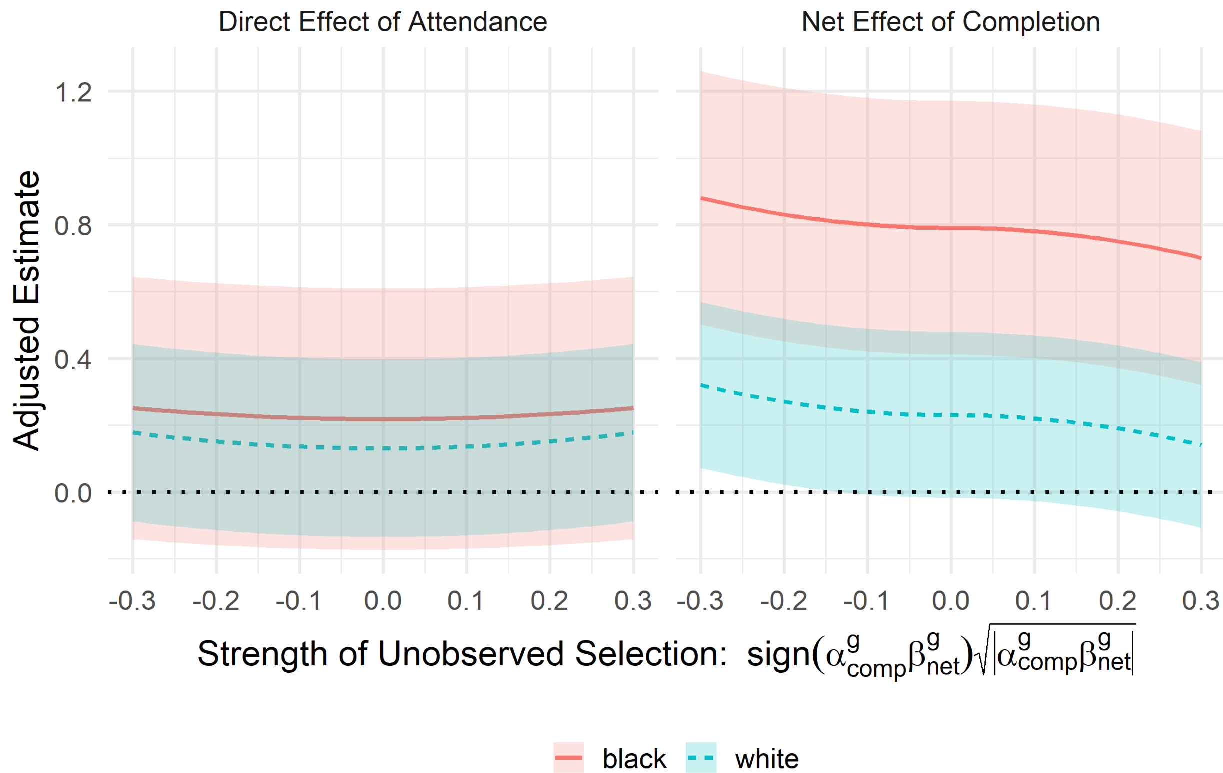

Figure H2: Sensitivity Results for the Direct Effect of Attendance (Δ^g_{att}) and the Net Effect of BA Completion (Δ^g_{comp}) on Men's Earnings

H.2 Bias Formulas

In this subsection, we derive bias formulas separately for Δ^g_{tot}, Δ^g_{att}, Δ^g_{net}. To simplify notation, we condition on $G = g$ implicitly throughout the exposition. First, let us consider a binary unobserved confounder U that affects both college attendance (A) and earnings (Y) and make the following simplifying assumptions: (a) $\mathbb{E}[Y|x, U = 1, a] - \mathbb{E}[Y|x, U = 0, a]$ does not depend on x and a; (b) $\Pr[U = 1|x, A = 1] - \Pr[U = 1|x, A = 0]$ does not depend on x (VanderWeele and Arah 2011). For $a = 0, 1$, we have

$$
\begin{aligned}
\mathbb{E}[Y(a)] &= \int \mathbb{E}[Y|x, u, a] dP(x, u) \\
&= \int \big(\mathbb{E}[Y|x, U = 1, a] \Pr[U = 1|x] + \mathbb{E}[Y|x, U = 0, a] \Pr[U = 0|x]\big) dP(x).
\end{aligned}
$$

Without adjusting for U, our estimator for $\mathbb{E}[Y(a)]$ will converge to

$$\begin{aligned}\mathbb{E}^*[Y(a)] &= \int \mathbb{E}[Y|x,a]dP(x) \\ &= \int \big(\mathbb{E}[Y|x,U=1,a]\Pr[U=1|x,a] + \mathbb{E}[Y|x,U=0,a]\Pr[U=0|x,a]\big)dP(x).\end{aligned}$$

Taking the difference between $\mathbb{E}^*[Y(a)]$ and $\mathbb{E}[Y(a)]$ yields

$$\text{bias}\big[\mathbb{E}[Y(a)]\big] = \int \big(\mathbb{E}[Y|x,U=1,a]-\mathbb{E}[Y|x,U=0,a]\big)\big(\Pr[U=1|x,a]-\Pr[U=1|x]\big)dP(x). \tag{17}$$

Substituting $a = 0,1$ into equation (17), taking the difference between $\text{bias}\big[\mathbb{E}[Y(1)]\big]$ and $\text{bias}\big[\mathbb{E}[Y(0)]\big]$, and applying assumptions (a) and (b), we obtain

$$\begin{aligned}\text{bias}[\Delta^g_{\text{tot}}] &= \underbrace{\big(\Pr[U=1|x,A=1]-\Pr[U=1|x,A=0]\big)}_{:=\alpha^g_{\text{tot}}}\underbrace{\big(\mathbb{E}[Y|x,U=1,a]-\mathbb{E}[Y|x,U=0,a]\big)}_{:=\beta^g_{\text{tot}}} \\ &= \alpha^g_{\text{tot}}\beta^g_{\text{tot}}.\end{aligned}$$

Next, consider a binary unobserved confounder U that affects both BA completion (M) and earnings (Y) but may itself be affected by college attendance (A). In addition, assume: (a) $\mathbb{E}[Y|x,a,z,U=1,m]-\mathbb{E}[Y|x,a,z,U=0,m]$ does not depend on x,a,z,m; (b) $\Pr[U=1|x,A=1,z,M=1]-\Pr[U=1|x,A=1,z,M=0]$ does not depend on x and z. For any a,m, we have

$$\begin{aligned}\mathbb{E}[Y(a,m)] &= \int \mathbb{E}[Y|x,a,z,u,m]dP(z,u|x,a)dP(x) \\ &= \int \big(\mathbb{E}[Y|x,a,z,U=1,m]\Pr[U=1|x,a,z]+ \\ &\qquad \mathbb{E}[Y|x,a,z,U=0,m]\Pr[U=0|x,a,z]\big)dP(z|a,x)dP(x).\end{aligned}$$

Without adjusting for U, our estimator for $\mathbb{E}[Y(a,m)]$ will converge to

$$\mathbb{E}^*[Y(a,m)] = \int \mathbb{E}[Y|x,a,z,m]dP(z|x,a)dP(x)$$

$$= \int \big(\mathbb{E}[Y|x,a,z,U=1,m]\,\Pr[U=1|x,a,z,m] \\ + \mathbb{E}[Y|x,a,z,U=0,m]\,\Pr[U=0|x,a,z,m]\big)dP(z|a,x)dP(x).$$

Taking the difference between $\mathbb{E}^*[Y(a,m)]$ and $\mathbb{E}[Y(a,m)]$ yields

$$\text{bias}\big[\mathbb{E}[Y(a,m)]\big] = \int \big(\mathbb{E}[Y|x,a,z,U=1,m] - \mathbb{E}[Y|x,a,z,U=0,m]\big) \\ \cdot \big(\Pr[U=1|x,a,z,m] - \Pr[U=1|x,a,z]\big)dP(z|a,x)dP(x). \tag{18}$$

Since $M=0$ when $A=0$, $\Pr[U=1|x,A=0,z,M=0] = \Pr[U=1|x,A=0,z]$. Therefore, $\text{bias}\big[\mathbb{E}[Y(0,0)]\big] = 0$. Substituting $a=1$ and $m=0$ into equation (18) and applying assumptions (a) and (b), we obtain

$$\begin{aligned}\text{bias}[\Delta^g_{\text{att}}] =& -\underbrace{\big(\Pr[U=1|x,A=1,z,M=1] - \Pr[U=1|x,A=1,z,M=0]\big)}_{:=\alpha^g_{\text{comp}}} \\ &\cdot \underbrace{\big(\mathbb{E}[Y|x,a,z,U=1,m] - \mathbb{E}[Y|x,a,z,U=0,m]\big)}_{:=\beta^g_{\text{net}}} \int \Pr[M=1|x,A=1,z]dP(z|A=1,x)dP(x) \\ =& -\alpha^g_{\text{comp}}\beta^g_{\text{net}} \int \Pr[M=1|x,A=1]dP(x) \\ =& -\pi^g_{\text{comp}}\alpha^g_{\text{comp}}\beta^g_{\text{net}}.\end{aligned}$$

Substituting $a=1$ and $m=0,1$ into equation (18), taking the difference between $\text{bias}\big[\mathbb{E}[Y(1,1)]\big]$ and $\text{bias}\big[\mathbb{E}[Y(1,0)]\big]$, and applying assumptions (a) and (b), we obtain

$$\begin{aligned}\text{bias}[\Delta^g_{\text{comp}}] =& \underbrace{\big(\Pr[U=1|x,A=1,z,M=1] - \Pr[U=1|x,A=1,z,M=0]\big)}_{:=\alpha^g_{\text{comp}}} \\ &\cdot \underbrace{\big(\mathbb{E}[Y|x,a,z,U=1,m] - \mathbb{E}[Y|x,a,z,U=0,m]\big)}_{:=\beta^g_{\text{net}}} \\ =& \alpha^g_{\text{comp}}\beta^g_{\text{net}}.\end{aligned}$$