SOCIAL SCIENCES

Heterogeneous returns to college over the life course

Siwei Cheng[1]*†, Jennie E. Brand[2]†, Xiang Zhou[3]†, Yu Xie[4]†‡, Michael Hout[1]†

College graduates earn higher wages than high school graduates by age 30. Among women, the advantages of a college degree decline somewhat as they age, although they are still substantial at age 50; for men, the advantage of a college degree grows throughout the life cycle. Most previous research on returns to higher education has focused on income at a single point in time or averaged over multiple years; our contribution is to study how returns vary by age. We also document how these patterns vary by the propensity of graduating from college. We find modest wage returns for mid-propensity college graduates, but large returns for low-propensity and, for men, high-propensity college graduates. Our results rely on propensity score–based matching combined with multilevel growth curve models applied to data from the National Longitudinal Survey of Youth 1979 cohort.

INTRODUCTION

Some people benefit more from a college education than others do. Most scholars agree that returns vary, but they debate both how much variation there is and why returns vary (*1*–*5*). The positive selection hypothesis states that individuals select into college on the basis of their anticipated payoffs to attending college, and those most likely to attend college reap the highest economic returns from college (*3*, *6*, *7*). By contrast, the negative selection hypothesis suggests that individuals who are unlikely to attend college benefit the most from a college degree. The latter hypothesis draws on a sociological tradition observing that college-going behavior is governed not only by rational choice but also by structural conditions, cultural and social norms and circumstances (*2*), social costs, and the active role of colleges and universities in choosing who will and will not attend (*8*). For those from advantaged backgrounds, attending college is a culturally expected outcome and less exclusively and intentionally linked to economic payoff than it is for those in less advantaged groups, for whom a college education is a novelty that may demand economic justification. This argument rests on a process of differential selection into college, described by several earlier studies (*2*, *4*, *9*). Some recent work expands on the differential selection hypothesis by noting that social scientists can only observe some of the factors that matter for selection into college (*10*, *11*). For example, differential unobserved selection may arise if socioeconomically disadvantaged students who go to college know things about themselves, say motivation or perseverance, that surveys cannot measure well but nonetheless affect their chances of graduating (*7*).

When adjudicating between the positive versus negative selection hypotheses, most previous research focuses on cross-sectional wages or average wages over several years. This literature is useful for understanding returns to college, especially in contexts where wages are stable across a career. However, decades of research suggest that wage inequality is produced gradually over the life course (*12*–*15*). According to human capital theory, high- and low-skill workers may accumulate their human capital over the life cycle at different paces, leading to their differential earnings trajectories (*16*–*19*). College education also entails categorical distinctions such that highly educated workers enjoy higher status in labor market institutions and occupy structurally advantageous positions with greater chances of upward mobility (*20*, *21*). Therefore, the labor market benefits of college generally emerge slowly over the career rather than instantaneously at any given point. A point-in-time or static measure of the college wage return may fail to capture this gradual process.

For our perspective, the conclusion about whether the returns to college are positively or negatively associated with the propensity to complete college, as well as the strength of this association, may depend critically on the life stage at which the outcomes are measured. Some people may reap the returns to college education early in their career, resulting in an immediate wage return upon entering the labor force, while others accumulate the economic returns to college more slowly and benefit more in the long run (*13*, *22*). If low-propensity individuals require explicit economic justification to go to college, they may be more strongly motivated to reap an immediate economic payoff (*23*). Conversely, high-propensity individuals may be less driven by an immediate economic return and more likely to choose careers with long-term wage growth. If so, their economic returns to college may accrue over time, resulting in more sustained wage growth. This reasoning is in line with a literature suggesting differences in future orientation and loss aversion by family background (*24*–*29*).

In this paper, we move beyond the cross-sectional perspective and examine the economic returns to college over the life course. Considering the relationship between college education and the life course wage trajectory, we posit that the economic return to completing a baccalaureate degree is contingent upon which years in life we assess (due to life cycle variation in returns) and which subpopulation we focus on (due to population heterogeneity in returns). Reduction of the life trajectory of the college wage return to a cross-sectional or average measure may obscure the long-term heterogeneous process through which the college return unfolds differently for different individuals.

Heterogeneous economic returns to college over the life course

Social scientists have long established that, other factors being equal, college graduates earn higher wages, have higher-status jobs, and are more likely to experience promotions and wage growth than those without a college degree (*30*–*32*). Although the literature has primarily focused on economic returns to a college degree at a given point in

[1]New York University, New York, NY, USA. [2]University of California, Los Angeles, Los Angeles, CA, USA. [3]Harvard University, Cambridge, MA, USA. [4]Princeton University, Princeton, NJ, USA.
*Corresponding author. Email: siwei.cheng@nyu.edu
†All authors contributed equally to this work.
‡Present address: Peking University, No. 5 Yiheyuan Road, Haidian District, Beijing 100871, P.R. China.

life, work over the last 20 years attends to the college wage premium over the life cycle (*12*, *13*, *17*, *19*, *33*). College graduates may invest more in their human capital through graduate education and on-the-job training (*34*), which contributes to their higher levels of job-specific human capital, greater chances of promotion, upward job mobility, and better job match quality (*16*, *17*, *34*, *35*). In addition, insofar as college increases skills, it may take time for employers to fully recognize and reward the skills employees acquired in college. These advantages, in turn, contribute to a divergence in wage trajectories by a college degree over the life course.

The sociological literature offers important additional insights into the structural and institutional mechanisms underlying college graduates' accumulation of economic advantage, over and above individuals' skills and credentials. A college degree marks a status distinction in the workplace, placing college graduates at structurally advantageous positions in the labor market that are associated with better career prospects and leading to a greater college wage premium during middle or late stages of their careers (*20*). For instance, a college degree can promote individuals' chance of overcoming institutional barriers surrounding occupations and moving into higher-earning occupations, which grants them better career prospects than their peers without college degrees. Furthermore, occupations that college graduates are likely to enter may yield occupational rents through institutionalized closure strategies such as licensing and credentialing (*36*). Within an organization, a college degree allows one to draw on the categorical distinction to claim resources and promotion opportunities, leading to a growing advantage over time (*21*).

Although past research has made notable progress toward understanding the dynamic, cumulative process of social stratification between individuals with and without a college degree, it has not examined potential heterogeneous effects of college on long-term wage trajectories. Individuals from different family backgrounds may differ not only in their likelihood of obtaining a college degree but also in the temporal pattern by which they reap the economic returns to college over the course of their careers. We summarize relevant literature below.

College as an equalizer

A college degree may serve as an equalizer that makes up for the socioeconomic disadvantage of individuals (*37*–*39*). Recent studies also suggest that crossing the barriers of college admission leads to large post-college earnings gains among academically marginal students, and that this effect is particularly strong among male and lower-income students, groups that are relatively unlikely to attend college (*2*, *5*). The earnings' effects of college admission for women and economically advantaged students were shown to be small and not statistically significant (*5*). Viewed from a life-course perspective, such an equalizing effect of college can help individuals from disadvantaged families not only secure rewarding economic positions in the labor market immediately upon college graduation but also circumvent the bleak career trajectories that they would have experienced had they not obtained a college degree (*40*). Thus, an equalizing effect of college suggests that the wage effect of college may be particularly strong among disadvantaged individuals and increasingly so as their career trajectories unfold.

The long reach of family background

On the other hand, family background may continue to shape individuals' participation in network-building and extracurricular activities during college (*41*, *42*) as well as their progress into advanced degrees after college graduation (*39*). These effects may well extend to the labor market. Individuals from advantaged family backgrounds may be favored by job recruiters (*43*) and are able to use social and cultural resources in their family networks (*44*), which, in turn, connect them to better job opportunities and prospects for wage growth. How these effects influence the long-term economic returns to college, however, depends on how they evolve over time: If the influence of parental background takes place mostly at early career stages and diminishes over time, then we expect to see a moderate reduction in the wage gap among college graduates by family background over the life course. If, however, the initial employment advantages of individuals from higher family backgrounds have a cumulative effect over the course of their careers, or if their parental social and economic resources continue to boost their careers at various life stages, we would expect to see an increase in the wage gap among college graduates by family background over the life course (*15*). This perspective is consistent with a recent study that finds that once selection processes are accounted for, the influence of parent income on child income in midlife is still substantial and about as strong among college graduates as among nongraduates (*45*).

Future orientation in career plans

How individuals anticipate the future, how they weigh future circumstances against present ones, and how they orient their behavior to the future can vary substantially by family background, leading to differences in career plans and earnings trajectories (*25*–*29*). Studies in sociology, psychology, and economics suggest that individuals from more advantaged family backgrounds place greater weight on long-term future returns, leading to a willingness to invest in further schooling or a career that has a steep wage trajectory and larger returns later in the life course (*26*). By contrast, for reasons ranging from social norms to financial burdens and family pressure, those from less privileged families place weight on generating more immediate returns. Unlike their more advantaged peers who have the luxury of exploration and "emerging" into adulthood, disadvantaged youth are on an "expedited path to adulthood," whereby they seek expediently to find a means to financial security (*23*). Family poverty and economic insecurity, as well as violence and uncertainty in longevity, lead to this orientation toward immediate economic rewards among disadvantaged youth. Although future and present time orientation is shaped in fundamental ways by childhood socioeconomic status (*29*), individuals from disadvantaged families who manage to attend and complete college may develop a stronger future orientation (*28*) and consequently experience an increasing wage profile relative to their noncollege graduate peers.

Interplay of work and family

Another important factor shaping an individual's trajectory in the labor market is the interplay of work and family transitions, particularly for women. While both a college degree and advantaged family background may boost individuals' earnings prospects in the long run, women may not be able to fully enjoy these benefits due to the career-impeding impact of child-bearing and child-rearing (*46*–*48*). Moreover, the negative impact of children on women's careers is particularly large among high-skilled women due to their high returns to work experience (*49*). These factors may not have an immediate effect during early life stages, at which most of the college-educated women may not yet have had their children. However, as the life cycle unfolds, these factors can greatly reduce the college premium for women around the time of child-bearing and child-rearing. We therefore expect a diminishing college premium among women during these years. Further, if the effects of college on women's marriage and fertility behaviors and their wages vary by family background (*50*, *51*), then the life cycle trajectories of the college premium may accordingly vary.

Unobserved selection

Last, the long-term economic returns to college may vary due to differential selection into college based on unobserved characteristics (*2*, *10*, *52*). Unobserved selection may be stronger among individuals from the most disadvantaged backgrounds. In this group, because of the high barriers to obtaining a college degree, those who complete college may be a particularly selective group who have a strong motivation for upward mobility, ambition for future achievement, and values that are closer to their peers from more advantaged families. As a result, what appears to be a large economic payoff to college among the most disadvantaged may partly reflect the unobserved selectivity of college graduates in this group. Likewise, the most advantaged individuals who do not complete college are also a selective group, and their selectivity out of college can influence the estimated college effect in this group. While the biases due to unobserved selection cannot be completely removed, how it influences the estimated causal effects of college will be gauged using a sensitivity analysis, assuming varying degrees of unobserved heterogeneity.

RESULTS

In our main analysis, we invoke an "unconfoundedness" or "selection on observables" assumption that conditional on a set of pre-college covariates, there are no additional confounders for the relationship between college graduation and earnings (*53*). Under unconfoundedness, causal effects can be estimated through the use of a propensity score (*53*). We estimate the propensity score for each individual in the sample as the probability of college completion by age 25 conditional on a set of observed covariates using a logit regression model. The goal is to obtain estimates of the propensity score that lead to adequate balance in precollege covariates that may affect both the likelihood of college completion and wages between the treated and control subsamples.

Using the estimated propensity score and matching methods described in Materials and Methods, we construct a sample in which all relevant covariates are roughly balanced between the treatment and control groups. Figure 1 compares the degree of covariate balance before and after matching for men and women, respectively. For each of the covariates included in the propensity score model, they show the absolute value of its standardized mean difference between college graduates and nongraduates before and after matching, calculated as its mean difference between the two groups divided by its SD among college graduates. We can see that in the matched samples, all of the pretreatment covariates have a standardized mean difference less than 0.25, a threshold suggested by current literature for assessing if a covariate is adequately balanced between treatment and control groups (*54*). As expected, the sizes of the treatment and control groups depend on propensity score. Among those with a propensity score < 0.02, 6 men and 3 women completed college by age 25, and 359 men and 577 women did not complete college; among those with a propensity score > 0.54, 200 men and 206 women completed college by age 25, and 82 men and 86 women did not. Given the small number of college graduates in the lower tail of the propensity

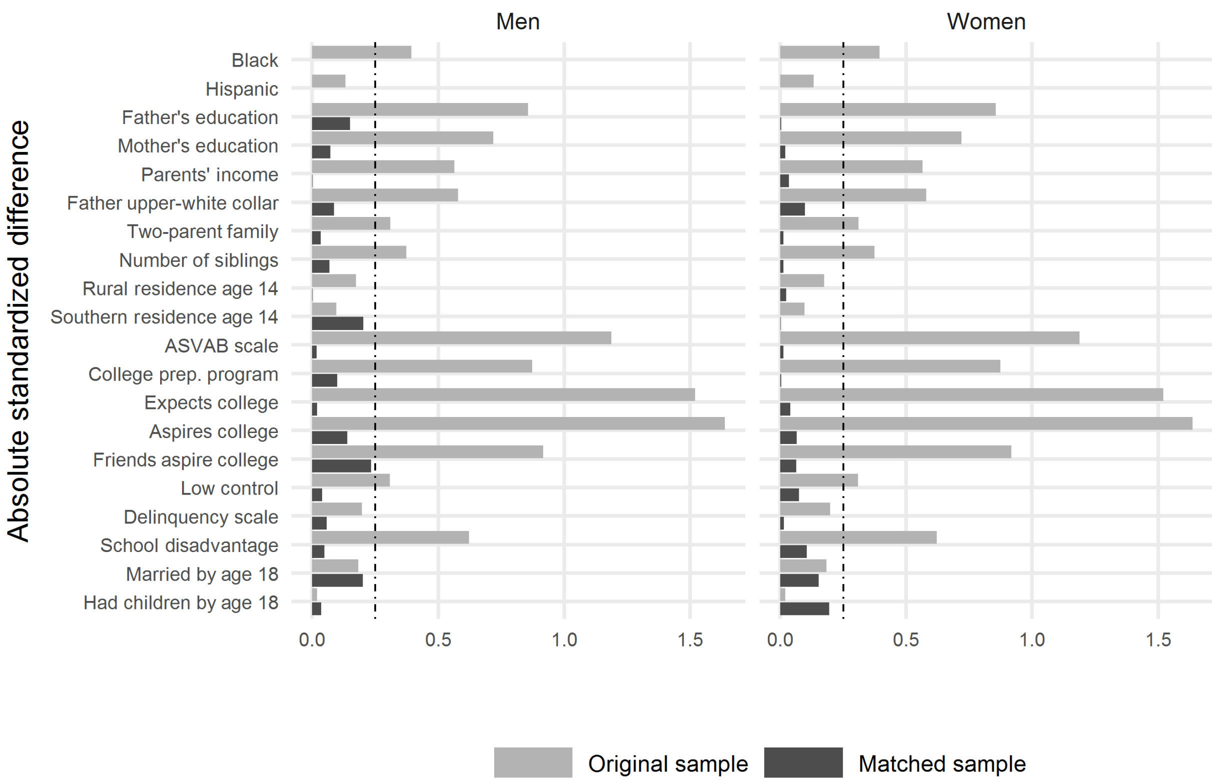

Fig. 1. Covariate balance before and after matching by gender. Note: For a given covariate, the absolute standardized difference is defined as the absolute value of its difference in mean between college graduates and nongraduates divided by its SD among college graduates. The vertical lines correspond to the value of 0.25.

score distribution, we conducted a robustness check by excluding from our analyses all individuals whose estimated propensity scores are in the bottom 20% of the propensity score distribution. The results are shown in fig. S2. Our main findings, particularly for men, are not driven by the lower tail of the propensity score distribution.

Figure 2 presents the histograms for the distribution of propensity score by gender and college completion. In our analytic sample, the average propensity score is 0.26 for men and 0.27 for women. The propensity score distribution differs greatly by college completion: The mean propensity score is estimated at 0.15 for men without a college degree, 0.54 for men with a college degree, 0.17 for women without a college degree, and 0.54 for women with a college degree. The propensity score distribution is concentrated at the lower end among noncollege graduates, and it is more dispersed among college graduates.

Figure 3 presents the estimated college wage premium across the propensity score values for different ages, based on the growth curve models specified in Materials and Methods. In keeping with the literature, we measure the college premium as the college/noncollege difference in log wage. The shaded band corresponds to 95% confidence intervals constructed from the variance-covariance matrix of the fixed effects components of the growth curve model. To highlight the range of the propensity score where the bulk of the population is concentrated, we show the propensity score on its percentile rank scale such that each grid along the horizontal axis represents the same (weighted) number of individuals. Figure S1 shows a parallel figure where the propensity score is on its original scale. Among men, the college premium exhibits a U-shaped pattern. The college premium ranges between approximately 0.4 to 0.7 at the lower end of the propensity score distribution, and it decreases as the propensity score increases, suggesting negative selection until reaching its lowest point around the median value of the propensity score. To the right of the lowest point, the college premium increases as the propensity score increases, suggesting positive selection. Hence, among men, the college wage premium is highest for those least and most likely to complete college, and lowest among individuals whose propensity score is in the middle. The magnitude of this heterogeneity is substantial: Evaluated at age 40, the lowest college premium among men is about 0.2, while the college premium can reach as high as 0.6 among the lowest-propensity individuals and 0.8 among the highest-propensity individuals. The confidence interval of the predicted college premium for the low-propensity levels is large and overlaps with that of the higher-propensity levels. This partly reflects the relatively small number of college graduates at the lower tail of the propensity score spectrum. Overall, the results provide suggestive evidence of a U-shaped pattern of returns, in line with recent literature using instrumental variable methods (*11*).

Among women, the college premium exhibits a mild L-shaped pattern at age 30: The college premium starts off high (nearly 0.4) at the lower end of the propensity score distribution, and decreases as the propensity increases until around 0.2 to 0.3. Unlike the pattern for men, in which the college premium rises thereafter, the college

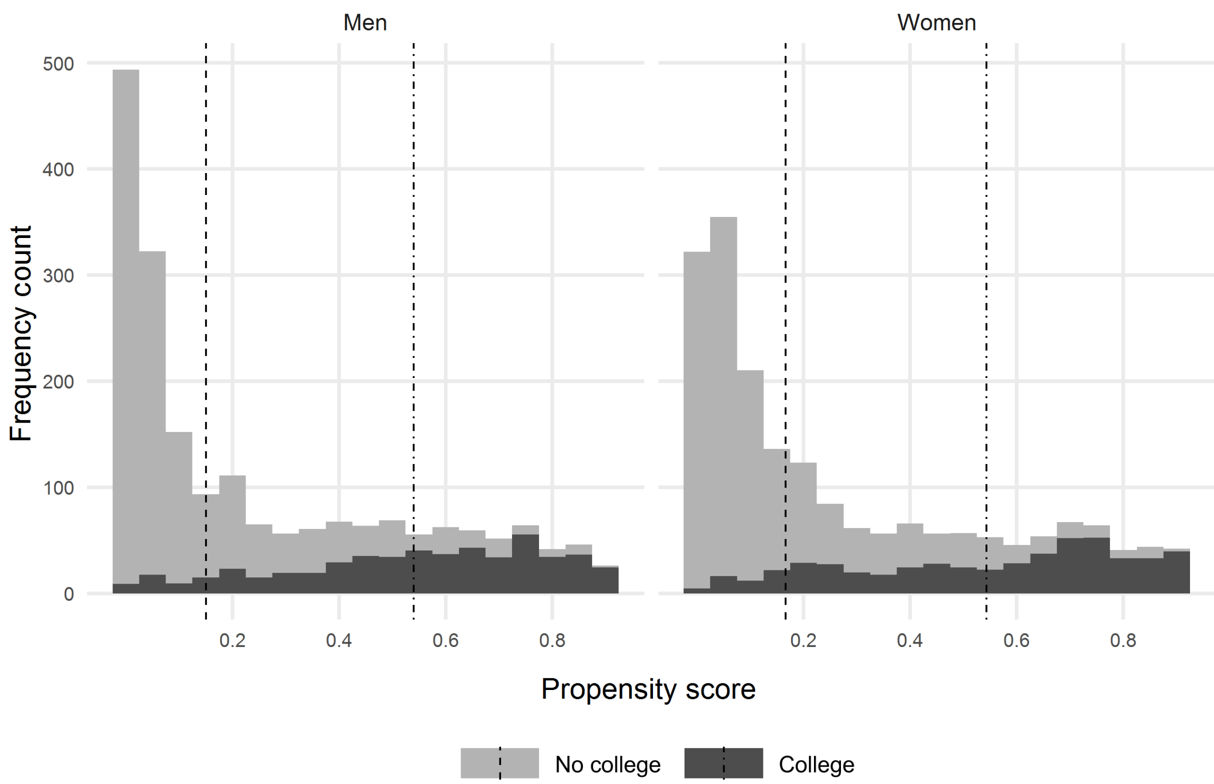

Fig. 2. Propensity score distribution by college completion and gender. Note: The histograms of the propensity score for noncollege graduates (light gray) and college graduates (dark gray) are stacked on top of each other. In the entire analytic sample, the average propensity score is 0.25 for men and 0.26 for women. In each plot, the dashed line shows the mean propensity score for nongraduates, and the dot-dash line shows the mean propensity score for college graduates. The mean propensity score is 0.15 for men without a college degree, 0.54 for men with a college degree, 0.17 for women without a college degree, and 0.54 for women with a college degree.

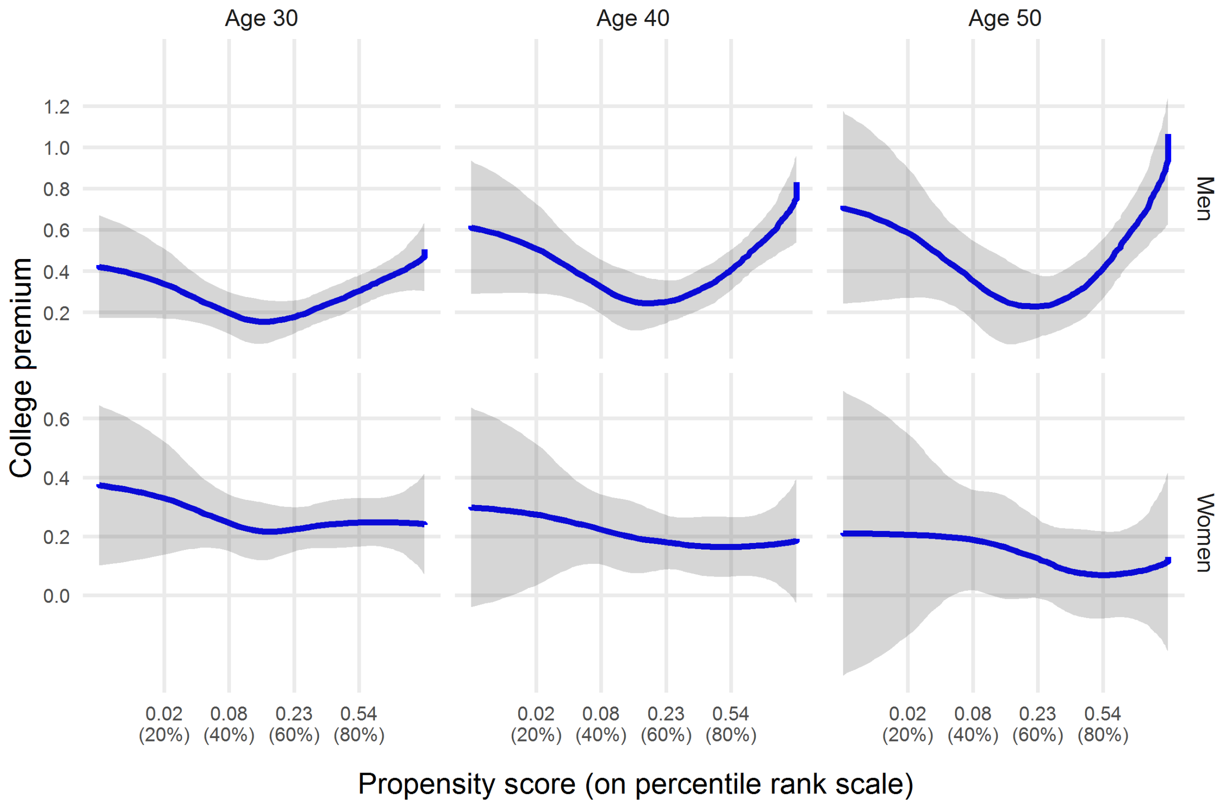

Fig. 3. Predicted college premium in log hourly wage by propensity score. Note: The horizontal axis shows the propensity score on its percentile rank scale such that each grid represents the same number of individuals. The labels of the axis ticks indicate the values of the propensity score at its 20th, 40th, 60th, and 80th percentiles. Error ranges correspond to 95% confidence intervals constructed from the variance-covariance matrix of the fixed effects components of the corresponding growth curve model.

premium among women changes little over the life course. We observe a moderate decrease in the college premium over age for women with highest propensity scores. This is consistent with our earlier discussion that the economic returns to college among women drop during the life course stages when demands for child-bearing and child-rearing are likely to be high for many women. To be sure, alternative measures of economic well-being, such as marital formation, assortative mating, or family income, may be more indicative of women's overall economic standing and should be considered in future work. For example, a study shows that high-propensity men and women have a larger effect of college on forming martial unions than do low-propensity men and women, for whom the effect of college on marital formation is negative (*51*). We focus on labor market earnings as the more direct and immediate indicator of economic well-being for both men and women.

To further examine the sources of variations in college returns observed above, we compute the age trajectories of predicted log hourly wage by college completion for individuals at low (20th percentile), middle (50th percentile), and high (80th percentile) values of the propensity score. The results are shown in Figs. 4 and 5 for men and women, respectively. These figures help illustrate the processes through which the college wage premium unfolded over the life course. Among men, across all propensity score levels, those with a college degree followed steeper wage trajectories than those without a college degree, resulting in a greater college premium at later life course stages. The large college premium among those with a low propensity to graduate from college is driven particularly by the stagnant and decreasing wage trajectories among men without a college degree, suggesting that a college degree helped these individuals circumvent bleak wage prospects. In addition, a college degree may have shaped their future orientation toward long-term economic gains over immediate economic returns. Among those with a high propensity score, the large college premium is driven mostly by college degree holders' higher growth rate over the life course, suggesting that a college degree afforded these individuals a cumulative advantage in wage attainment that unfolded over their lives.

The trajectories for women tell a rather different story. At all propensity score levels, the predicted wages among women with a college degree are higher than those among women without a college degree, but the two trajectories converge, instead of diverge, with age. This results in a moderate reduction in the college premium among women over the life course. One possible explanation is that women suffer a wage penalty when they become mothers, and the motherhood penalty may be particularly large among high-skilled women (*49*). The estimated college wage premium is largest among women who are least likely to complete college, suggesting a "negative selection" story.

As a measure of the likelihood of completing college given observed covariates, the propensity score can be seen as a one-dimensional summary of individual characteristics. We now zoom in on key

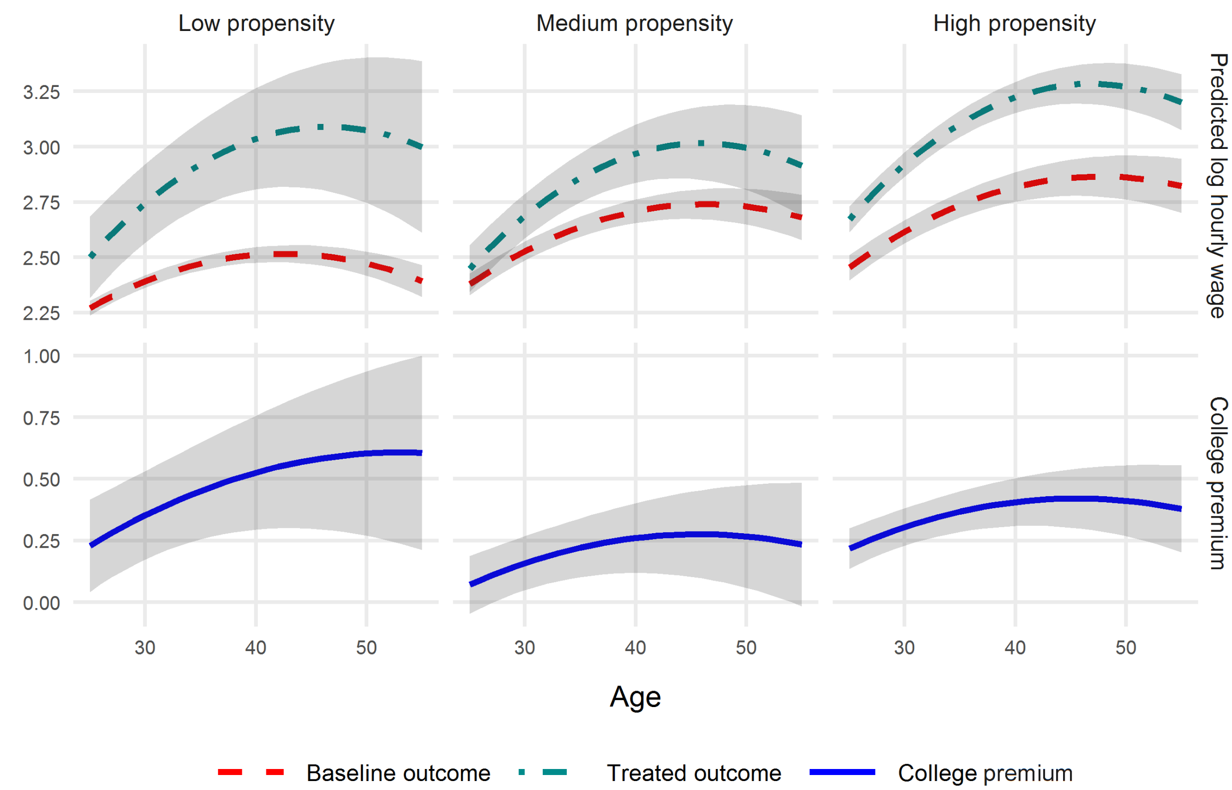

Fig. 4. Predicted age trajectories of log hourly wage with and without a college degree at different propensity score levels, men. Note: The left, middle, and right panels correspond to predicted age-wage profiles at the 20th, 50th, and 80th percentile of the propensity score. The upper and lower panels show predicted log hourly wage and predicted college premium, respectively. Error ranges correspond to 95% confidence intervals constructed from the variance-covariance matrix of the fixed effects components of the growth curve model.

components of the propensity score and examine factors representing social background and achievement that might play different roles in shaping the college wage premium: mother's education (Fig. 6), parental income (Fig. 7), and cognitive ability (Fig. 8). To purge the influence of parental income and mother's education on our measure of cognitive ability (the ASVAB test score), we fit a linear regression of cognitive ability on mother's education and terciles of parental income separately for men and women and then use the residuals from these regressions to define cognitive ability groups. In short, at the more detailed level of components constituting the propensity score summary measure, we focus on one key covariate at a time and ask whether, for example, low income students would benefit from a bachelor's degree more than higher income students, while low ability students would not benefit more than higher ability students. As in our propensity score–based analyses, we fit growth-curve models similar to Eqs. 2 to 4 in Materials and Methods except that the propensity score splines are replaced with group indicators for mother's education, parental income, and (residualized) cognitive ability.

The variation in the college premium by mother's education is largely consistent with the U-shaped pattern for men and the L-shaped pattern for women from the propensity score–based analyses. The variation in the college premium by parental income exhibits a pattern of negative selection, which is more pronounced at younger ages for men and at older ages for women, suggesting that the effect of college on wages is greatest among individuals from low-income families. Last, the variations in the college premium by cognitive ability exhibit a U-shaped pattern for both men and women. Thus, the wage returns to college for men and women were both higher among low- and high-ability individuals. Together, these findings suggest that individuals from the most disadvantaged backgrounds—whether it is measured using propensity of college, mother's education, parental income, or cognitive ability—received higher economic returns from college than their peers from the middle of the socioeconomic spectrum.

A sensitivity analysis on unobserved selection

Throughout our previous analyses, we have assumed that the pretreatment covariates included in our propensity score model capture all relevant factors that may confound the causal relationship between college completion and earnings. This assumption is strong, untestable, and unlikely to hold true in reality. Given that unobserved selection can bias our estimated patterns of heterogeneous returns to college, we now conduct a formal sensitivity analysis to explore the direction and magnitude of potential bias. The methodological details are described in Materials and Methods. The results are shown in Fig. 9. We can see that for men, the U-shaped pattern of heterogeneous returns to college is quite robust. For women, the estimated magnitude and patterns of college premium are more sensitive to unobserved selection, but the moderate L-shaped pattern still roughly holds under various assumptions on the strength of unobserved selection.

In general, the direction of the bias depends on the pattern of potential unobserved selection. For example, the argument for "ability

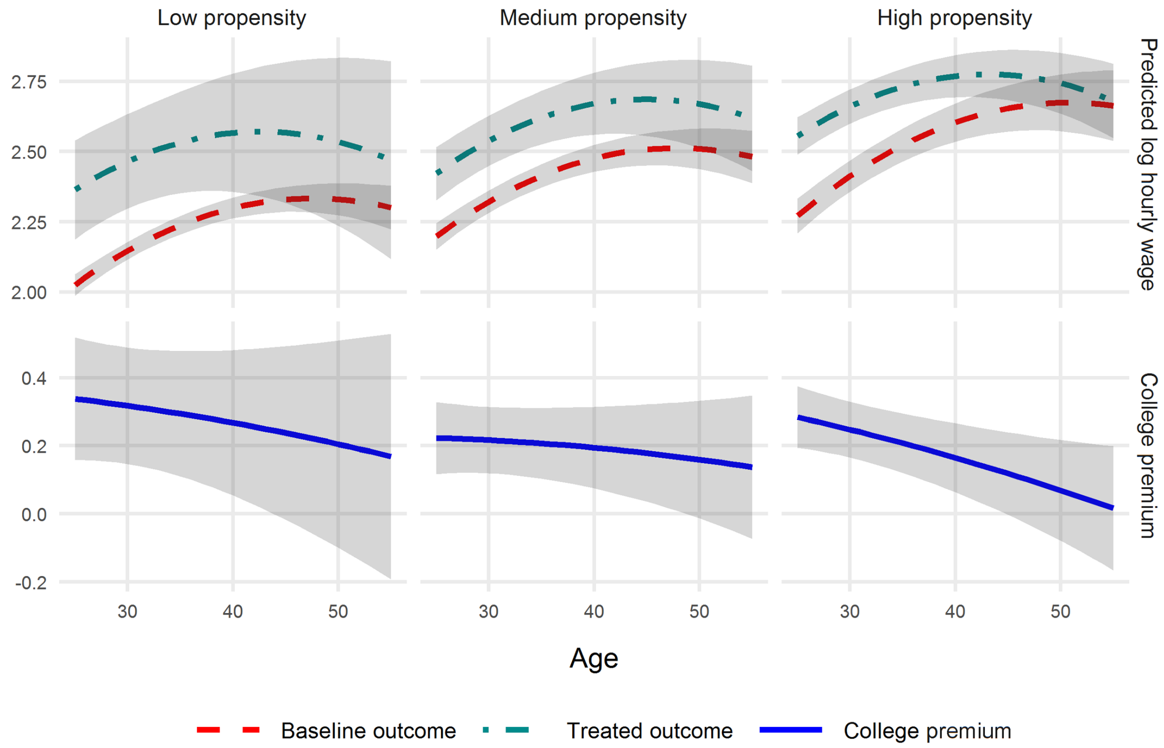

Fig. 5. Predicted age trajectories of log hourly wage with and without a college degree at different propensity score levels, women. Note: The left, middle, and right panels correspond to predicted age-wage profiles at the 20th, 50th, and 80th percentile of the propensity score. The upper and lower panels show predicted log hourly wage and predicted college premium, respectively. Error ranges correspond to 95% confidence intervals constructed from the variance-covariance matrix of the fixed effects components of the growth curve model.

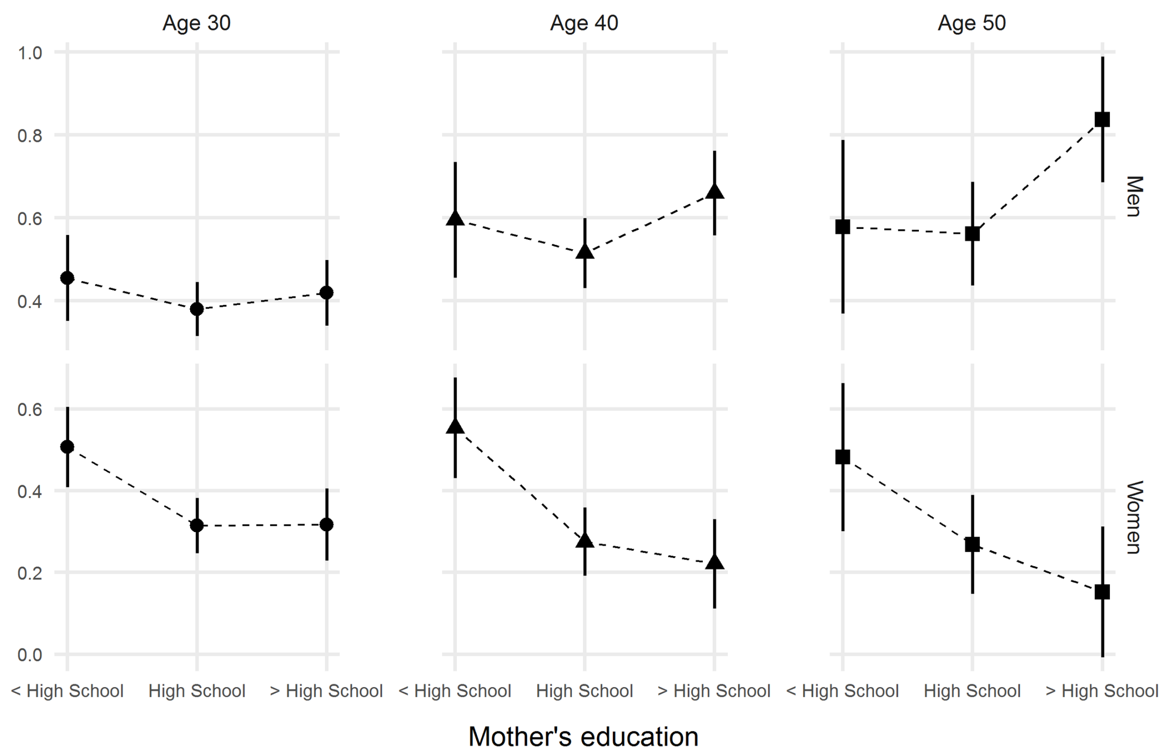

Fig. 6. Predicted college premium in log hourly wage by mother's education. Note: Error ranges correspond to 95% confidence intervals constructed from the variance-covariance matrix of the fixed effects components of the corresponding growth curve model.

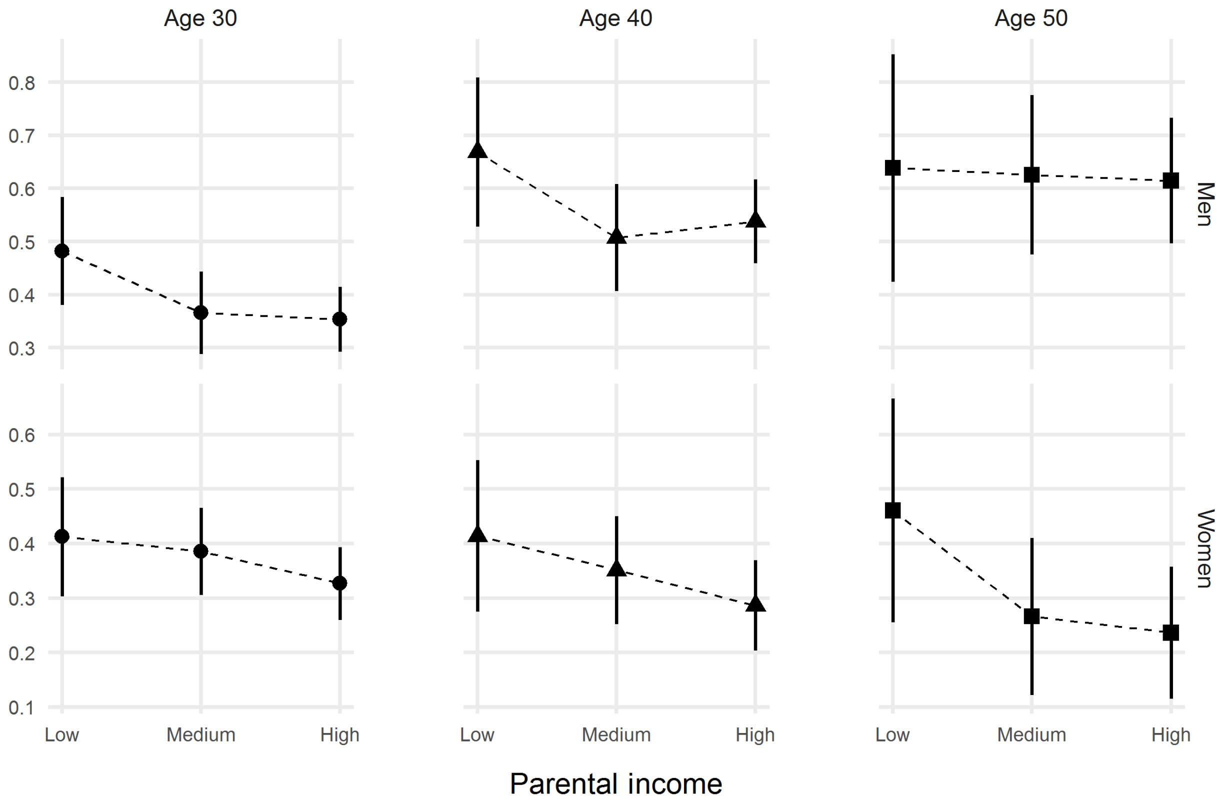

Fig. 7. Predicted college premium in log hourly wage by parental income. Note: Low, medium, and high levels of parental income correspond to the first, second, and third terciles of the income distribution. Error ranges correspond to 95% confidence intervals constructed from the variance-covariance matrix of the fixed effects components of the corresponding growth curve model.

bias" (*55*) predicts a negative association between the error terms ϵ and V defined in Eqs. 5 and 7, that is, more capable individuals tend to be associated with both more education and higher overall earnings. In this case, the estimated causal effects of college would suffer from an upward bias. On the other hand, more recent research considering individual heterogeneity in both baseline earnings and the causal effect of college has found support for the "comparative advantage" argument, which implies negative sorting on level ($\rho_{\epsilon V} > 0$) and positive sorting on gain ($\rho_{\eta V} < 0$) (*3*, *7*). In other words, it is hypothesized that individuals who actually completed college would be worse off than those who did not if both groups had not completed college, although the former group benefits more from college education than the latter group would had they completed college. In this case, our estimated causal effects of college would be downwardly biased, and the U-shaped pattern of heterogeneous returns would be even more pronounced than those observed in the main analyses (see the lines corresponding to a positive $\rho_{\epsilon V}$).

DISCUSSION

Do some people get more out of college than others do? Interest in heterogeneous college returns along various dimensions has a long history. While the debate over who benefits most from college continues, over the last two decades, several studies have shown larger economic returns for black than for white students (*56*, *57*) and larger returns for low-income than for high-income students (*56*, *58*). Others have suggested larger returns for students on the margin of school continuation as identified by various instrumental variables (*32*, *59*) [see (*8*) for review]. Research also suggests that the returns to college vary by gender given the cost of childbearing to women's career progression and wage attainment (*46*, *49*, *60*–*62*) and the selectivity of women's labor force participation (*63*). Because so many factors influenced who went to college and who, among those who went, earned a degree—factors such as peers' influence, school environment, proximity to a college or university, and local labor market conditions (*64*)—a decade ago researchers shifted their attention away from specific characteristics and began to use summary measures of the propensity to graduate from college in the search for heterogeneity (*2*, *3*, *10*, *50*). Individual-level variability in the causal effect of college on economic outcomes can shed light on the various explanations as to why some individuals complete college, while others do not, as well as on how college premium might slow social mobility (*65*) and contribute to social stratification (*66*).

Existing literature in both economics and sociology on the heterogeneous effects of college on economic outcomes has, so far, almost exclusively focused on cross-sectional measures of economic outcomes at certain ages or averaged across the lifetime. Our contribution lies in considering variation in college returns over the life cycle. The full labor market benefits of a college degree emerge gradually over a career. Hence, college may be associated with not only a higher initial wage but also a faster rate of wage growth. Low-propensity college graduates may seek more immediate rewards, while high-propensity college graduates might take a long-term orientation. On the other hand, a college degree may help low-propensity

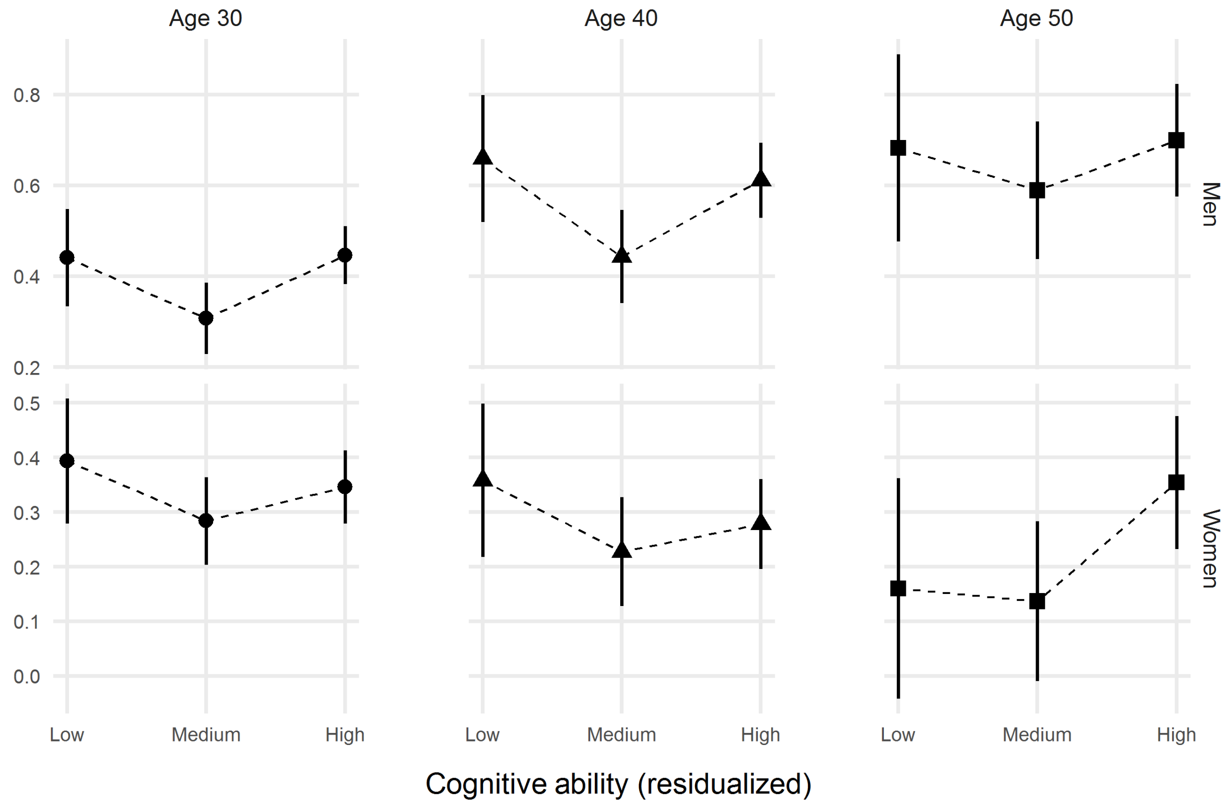

Fig. 8. Predicted college premium in log hourly wage by cognitive ability. Note: Low, medium, and high levels of cognitive ability correspond to the first, second, and third terciles of the distribution of the residualized ability scores. Error ranges correspond to 95% confidence intervals constructed from the variance-covariance matrix of the fixed effects components of the corresponding growth curve model.

individuals avoid bleak career prospects that they would have experienced without college. The varying returns to college degrees by propensity of college might emerge gradually and cumulatively over their life course. That is, we should expect heterogeneity in the trajectories of college returns. We expected this heterogeneity to be associated not only with young people's observed attributes but also with unobserved characteristics that affect selection into college.

Using a causal inference framework, we analyzed longitudinal wage trajectories with data from the National Longitudinal Survey of Youth 1979 cohort, with propensity score–based matching and multilevel growth curve models. In particular, we examined the variability in the college effects throughout the life course by the estimated propensity of completing a bachelor's degree. Men's wage return to college by propensity scores declines from the lowest propensity to the median propensity and then rises, forming a U-shaped pattern. This pattern becomes more pronounced at older ages. Together, these findings indicate that a college degree brought long-term and cumulative economic returns to men who were least and most likely to complete college. The U-shaped pattern of returns by propensity becomes more pronounced with age. By age 50, a college degree helped the most disadvantaged individuals avoid bleak labor market prospects that they would have experienced without college; meanwhile, it boosted the economic gain of the most advantaged individuals by placing them on a steeper wage trajectory than their noncollege counterparts.

Women's wage return to college at age 30 is highest for the lowest-propensity women, declines through the median propensity, and stays at that level (0.2) through the whole range of above-median propensities. Women's wage returns do not increase with age as men's do; they grew more uncertain and decline with age. This finding is consistent with previous research showing that women, particularly highly educated, high-skilled women, suffer a motherhood wage penalty relative to their nonmother peers (*49*, *62*). Men's and women's early returns to college graduation are similar except at high propensities; above the median propensity, men's returns surpass women's, and men's excess return grows with propensity score. Men's returns to college graduation are consistently higher than women's at age 40, more so at higher than lower propensities. Men's returns rose and women's probably fell between ages 40 and 50 so that the gender gap at each propensity score increased. Wide confidence intervals prevent a definitive conclusion, but men's point estimates are consistently higher than women's at age 50.

In conducting the analysis and reaching the above conclusion, we incorporated a number of methodological considerations to ensure that our results are robust. Specifically, we used a flexible specification of the propensity score model to allow for nonlinear and interaction effects of the predictors of college completion, as well as a flexible specification for the relationship between the college wage return and the propensity of college. By doing so, we are able to demonstrate the variations in the economic returns to college across the entire propensity score distribution. Unlike previous work that has

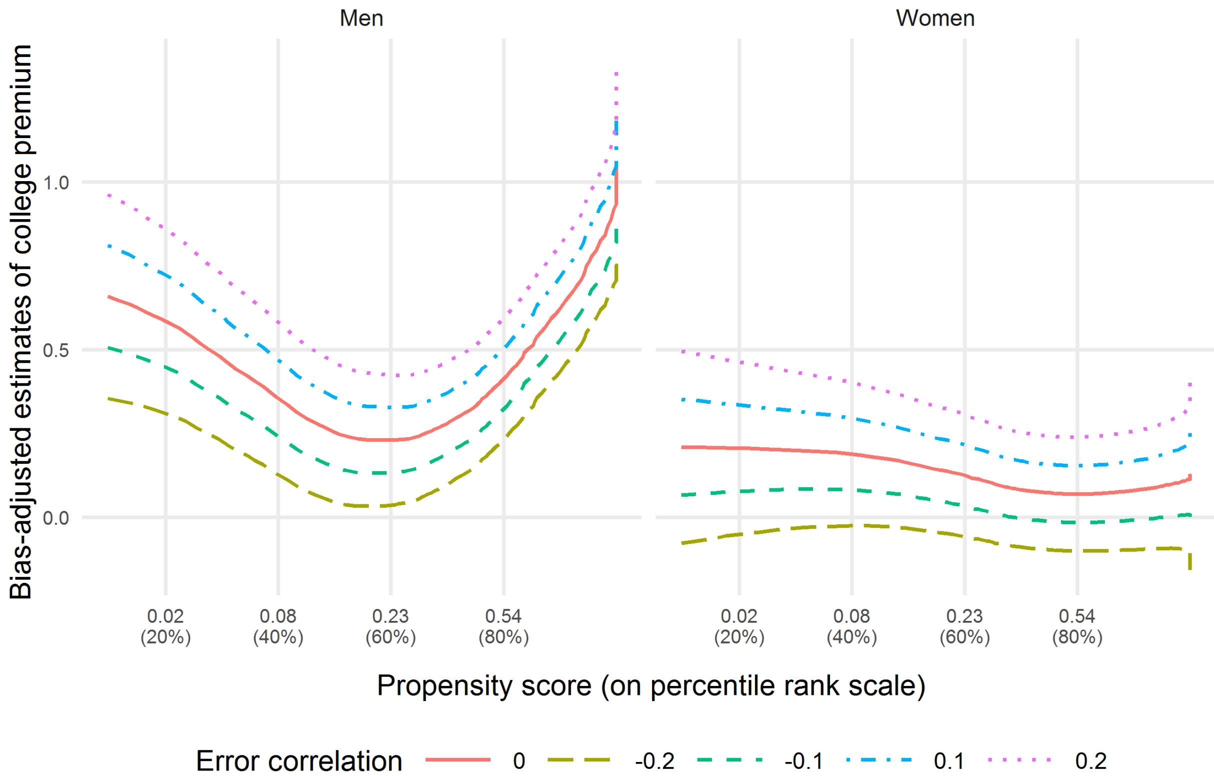

Fig. 9. Bias-adjusted estimates of college returns at age 50 by gender. Note: The horizontal axis shows the propensity score on its percentile rank scale such that each grid represents the same number of individuals. The labels of the axis ticks indicate the values of the propensity score at its 20th, 40th, 60th, and 80th percentiles.

usually assumed a monotonic pattern of college economic returns, our flexible approach has revealed a nonlinear relationship between the wage returns to college and the propensity to complete college. Results differ for men and women. Given the possibility that selection into college completion may involve individual traits that are unobserved in our data, we have also conducted sensitivity analysis to attend to potential unobserved selection into college and its implications for the college wage premium. Although it is not possible to directly address the problem of unobserved selection with observed data, we have shown that the patterns of heterogeneous returns to college for men and women are robust across a reasonable range of selection-related parameters.

On a broader level, our results illustrate the value of adopting a longitudinal perspective to understand the economic returns to college. With the rise of economic inequality and growing concerns about disparities in socioeconomic opportunity in society, the value of a college degree continues to stand in the center of debates among scholars, policymakers, and the general public. We show that when considering how college affects economic attainment, it is important to compare individuals' life course trajectories to their "counterfactual trajectories" that they would have followed if they had not completed college. While disadvantaged college graduates may still lag behind their more advantaged peers in economic outcomes, they are nevertheless significantly better off than their counterfactual trajectories if they had not completed college. A college degree helps them circumvent the bleak wage trajectories they would have experienced without a college degree. While this benefit may seem moderate at the beginning, it grows larger and larger over a career for men. We wonder whether socially disadvantaged youth at college-attending ages are as knowledgeable as their more advantaged peers about the college premium pattern over the course of a career. If not, knowing this fact per se may motivate more of them to attend and complete college.

College attendance and completion involves a complicated mix of choices made and constraints faced by students and colleges (*67*). Students and colleges that assume that only the "top" students benefit from college might rethink these assumptions in light of our results. College pays off across the whole spectrum of preparation (*68*). Colleges and universities that reject most applicants could reach below current cutoff points to admit more nontraditional students, confident that they, too, will benefit. Nontraditional students have lower completion rates in large part because they are "underplaced" (*68*, *69*). Our results over the life course suggest that the economic benefits of enrolling traditionally disadvantaged students are long-lasting; among women, they persist through age 50, and among men, they increase over the career. Our findings add another dimension to understanding the persistence of class-based inequality in college education.

MATERIALS AND METHODS

Data

We use data from the 1979–2014 waves of the National Longitudinal Survey of Youth 1979 cohort. This nationally representative longitudinal dataset provides rich information on respondents' sociodemographic characteristics, family background, achievement, skills, educational attainment, and long-term wage trajectories from early to late career. Our sample is restricted to individuals who were 14 to 17 years old at the baseline survey in 1979 ($n = 5582$) and who later completed at least the 12th grade ($n = 4548$). These sample restrictions are set to ensure that all variables used to predict college are measured before college, particularly ability, and to compare college graduates to those who had completed at least a high school education. The treatment variable is whether a respondent completed college by age 25. About one-quarter of the sample completed college by age 25 (*70*).

College experience differs by social background, resulting in "treatment heterogeneity." About one-third of those who did not complete college by age 25 had attended some college (one-fourth attend a 4-year college). This ranges from about one-fourth of low-propensity noncollege graduates to three-fourths of high-propensity noncollege graduates (one-fifth attended a 4-year school among the low propensity, and two-thirds did so among the high propensity). Fewer than 5% of noncollege graduates by age 25 went on to complete college within the next 5 years—which ranges from 3% among the low propensity to 12% among the high propensity. Almost all high-propensity college graduates attended a 4-year college by age 20 but so did over 85% of low-propensity college graduates. Among those who completed college, about 10% attended a highly selective school (according to 1980 Barron's Profiles), which ranges from 3% among low-propensity college graduates to 15% among high-propensity graduates. About one-third of college graduates went on to obtain a graduate degree, with a range of 26% among low-propensity to 38% among high-propensity college graduates.

We use inflation-adjusted log hourly wage of the person's current or primary job as the outcome variable. We measure hourly wage as a time-varying variable and obtain the age profiles for the respondents from age 25 to 55. We code zero wages as missing and use the person-specific wage trajectories estimated from the growth curve model to predict potential wages for person-years with a zero or missing wage. We use model-based prediction instead of zero (or a close-to-zero constant such as \$1). As the employment selection literature suggests, potential wages in the labor market for those without a positive wage are likely somewhere below their reservation wage, rather than zero (*63*, *71*, *72*).

Analytic strategy

Our main empirical analysis proceeds in three steps. In the first step, we use a large array of individual-, family-, and school-level covariates to estimate the propensity of completing college. In the second step, using the propensity scores constructed in the previous step as a measure of closeness, we construct a matched sample where each unit in the treated group is matched to the closest units in the control group, and each unit in the control group is matched to the closest units in the treatment group. This matching procedure results in a sample for estimating the average treatment effect (ATE). In the third step, we estimate multilevel growth curve models on the matched sample and use the model estimates to predict the trajectories of the college wage premium over the life course.

Estimating the propensity of college completion

To estimate the propensity of college completion, we adopt an iterative procedure outlined by (*53*), which leads to a fairly flexible specification with good balancing properties. The procedure is meant to improve upon common specifications that involve simply including all covariates additively. In principle, one could also run the models for men and women separately. However, a separate model assumes that every covariate has an interaction with gender. We ran every possible interaction with gender, and very few of the interaction terms were statistically significant. Hence, only a couple were retained in the final model. Having a very flexible model specification enables us to take the gender interactions, and all possible interactions, into consideration.

In the first step, we begin with a baseline theoretically motivated set of covariates K_B, which includes covariates that are a priori viewed as important for explaining the treatment and plausibly also relate to the outcome. Our basic covariates include those used in several papers on college effects by Heckman and colleagues [e.g., (*73*, *74*)]: gender, race, mother's education, fathers' education, family income in 1979, whether the respondent grew up with both biological parents, number of siblings, rural and southern residence, and cognitive ability as measured by the 1980 Armed Services Vocational Aptitude Battery (ASVAB), adjusted for age and standardized following earlier work (*75*). These variables are theoretically important predictors of selection into college and are included in most models of college attainment; they remain in the propensity model throughout our covariate selection procedure.

The second step is to consider a number of additional possible covariates in turn, including measures of the following: family background (i.e., father's education, father's occupation, mother's occupation, knowledge of father's education, and religious affiliation), achievement and psychosocial skills (i.e., whether the student was enrolled in a college-preparatory curriculum in high school, educational aspirations, educational expectations, Rotter locus of control, Rosenberg self-esteem, and a scale of delinquent activity), school characteristics (i.e., percentage of students classified as disadvantaged and percentage of students classified as black or Hispanic), and family formation (i.e., a scale of traditional family attitudes, marital status at age 18, and had a child by age 18). We also considered additional coursework and rank in high school class variables, but these indicators suffered from many missing values and were not especially predictive beyond the factors already included in the model. This is an iterative process where, in each step, we decide whether to include an additional covariate based on a set of logistic regression models. We add the covariate with the largest likelihood ratio statistic that exceeds a preset constant (1.0, or a z-statistic of 1.0), and the process repeats. This step involved 176 logistic regressions and a resulting model with 22 covariates. Some covariates, including mother's occupation, knowledge of father's education, religion, school racial composition, self-esteem, and family attitudes, did not reach the threshold for model improvement.

In the third step, we decide which of all possible higher-order and interaction terms to include in the model. We follow the same procedure as above, where we add the term with the largest likelihood ratio statistic that exceeds the preset constant (in this case, 2.71, or a z-statistic of 1.645). This procedure involved 3527 regressions. The resulting model includes 22 linear terms, 1 higher-order term (mothers' education squared), and 11 interaction terms. The resulting terms include a squared term for mother's education and interactions between family size and southern residence, gender and educational aspirations, parental income and delinquency, mother's education and educational aspirations, parental income and educational expectations, Hispanic and southern residence, ability and college preparatory program, educational expectations and aspirations, father's occupation and college preparatory program, black and college preparatory program, and rural and college preparatory program. Last, we eliminate units with no common support. We restrict our analyses to the region of common support by trimming the tails of the propensity score distribution. We eliminate 443 noncollege graduates with a very low propensity score ($P \leq 0.004$) and 20 college graduates with a very high propensity score ($P \geq 0.923$). Among our remaining sample, the college and noncollege groups have common support at both low (<0.1) and high (>0.9) levels of the propensity score. This ensures that we have nonmissing observations to estimate college wage returns across the propensity score distribution.

Matching

After estimating the propensity scores, we use exact matching on race and ethnicity, combined with nearest-neighbor matching on the linear propensity score, i.e., logit $[\hat{p}]$, to construct two matched samples, one for men and one for women. In matching, the linear propensity score is a preferred metric to the raw propensity score because the former does not penalize differences in pretreatment covariates at the tails of the propensity score distribution (*53*). For example, on the raw propensity score scale, a treated unit with $\hat{p} = 0.1$ is considered as close to a control unit with $\hat{p} = 0.15$ as to a control unit with $\hat{p} = 0.05$. However, in terms of the covariates, the treated unit tends to be much closer to the former than to the latter. The linear propensity score, by transforming $\hat{p}$ back to the scale of the covariates, does not suffer from this distortion.

We use one-to-five matching with replacement. That is, for each college graduate (non-graduate), we find five nongraduates (college graduates) with the same race and ethnicity and the closest linear propensity scores, and include them in the matched data with evenly distributed fractional weights. Specifically, each matched nongraduate (college graduate) is assigned a weight equal to $w_i/5$, where w_i is the NLSY (National Longitudinal Survey of Youth) sampling weight of the college graduate (nongraduate) to which the unit is matched. Thus, in the matched dataset, the same unit may appear multiple times, either as the primary unit of interest or as a match for a unit with the opposite treatment status. To simplify analysis, we collapse multiple records of the same unit onto a single record by summing up the weights. The final matched sample consists of 2079 men (423 college graduates and 1656 nongraduates) and 2045 women (432 college graduates and 1613 nongraduates).

Growth curve models

The matched sample created above represents a pseudo-population in which, under the assumption of unconfoundedness, treatment status is orthogonal to the propensity score as well as covariates that are predictive of college completion. Another important advantage of adopting this matching method is that, because the matching units are all constructed on the person level, information on within-individual year-to-year wage linkages is preserved in this matched sample, and this facilitates our modeling of person-specific wage trajectories.

Next, to analyze wage trajectories for individuals with and without a college degree, we estimate a growth curve model for log hourly wage (W_{it}) with our matched sample. The model contains two levels. At level 1, log hourly wage is specified as a function of age minus 30 (t):

$$W_{it} = \beta_{0i} + \beta_{1i} t + \beta_{2i} t^2 + e_{it} \tag{1}$$

Here, β_{0i}, β_{1i}, and β_{2i} represent the person-specific intercept, slope, and the coefficient on quadratic time, respectively. We then predict these coefficients using the specifications below

$$\beta_{0i} = \gamma_{00} + \gamma_{01} D_{\text{college}} + \Sigma_{k=1}^{3} \gamma_{02,k} S_k + \Sigma_{k=1}^{3} \gamma_{03,k} D_{\text{college}} \cdot S_k + u_{0i}, \tag{2}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11} D_{\text{college}} + \Sigma_{k=1}^{3} \gamma_{12,k} S_k + \Sigma_{k=1}^{3} \gamma_{13,k} D_{\text{college}} \cdot S_k + u_{1i} \tag{3}$$

$$\beta_{2i} = \gamma_{20} + \gamma_{21} D_{\text{college}} \tag{4}$$

In the above equations, S_1, S_2, and S_3 constitute a basis for a natural cubic spline of the estimated propensity score, where two knots are chosen at the tertiles of the propensity score distribution. Natural cubic splines are a flexible yet parsimonious tool for modeling nonlinear relationships [(*76*), pp. 144–146]. Alternative specifications on the number and positions of knots yield substantively identical results. Note that the interaction terms between these cubic splines and treatment status (D_{college}) capture potentially heterogeneous treatment effects on both the initial wage and the rate of wage growth by the propensity score. The cubic spline specification has the advantage over a linear propensity score specification in that it accounts for nonlinearity in predicted wages and the treatment effect; it is also more efficient and less arbitrary than grouping propensity scores into bins. We interact the cubic splines with treatment status for the random intercept and random slope.

There are two major benefits of using a growth curve model rather than directly calculating the average wage at each age. First, while descriptive statistics on average wages provide aggregate patterns in the population, the growth curve model makes more efficient use of the within-person linkages in our longitudinal data. Second, because wages are only observed among individuals who are working at the time of interview, the observed wage trajectories among nonzero wage earners may not be representative of the population at a given age, and this employment selection process may also vary by age. The growth curve model estimated on individual-level data, instead, allows us to extrapolate the trajectories in observed wage-earning years to years in which the person is not employed. Therefore, it alleviates the problem of employment selection. We estimate the growth curve models using the lme4 package in R (*77*). We construct confidence intervals from the estimated variance-covariance matrix of the fixed effects components of the growth curve model, which should be seen only as an approximation to the true confidence intervals that account additionally for the uncertainty associated with propensity score estimation and matching.

We further test the hypothesis that the treatment effect does not vary by the propensity score among men and women. Specifically, we ran an additional growth curve model without the two-way interaction effects between college completion and propensity score splines and the three-way interaction effects between college completion, propensity score splines, and age and conducted a likelihood ratio test by comparing the models with and without these interactions. Our results suggest that the null hypothesis of no effect heterogeneity can be safely rejected for men (chi-squared = 18.66 with 6 degrees of freedom, $P = 0.0047$), but not for women (chi-squared = 2.34 with 6 degrees of freedom, $P = 0.8854$).

As a robustness check, we implemented a regression-imputation (RI) approach and a doubly robust approach (DR), which involve specifying an outcome model given all pretreatment covariates and imputing potential outcomes for individuals within different propensity score strata. The findings, detailed in the Supplementary Materials, are highly consistent with our main results.

A sensitivity analysis on unobserved selection

Our approach to assessing heterogeneous returns to college by the propensity score can be seen as a special case of the marginal treatment effect (MTE) approach to assessing treatment effect heterogeneity under a generalized Roy model (*3*,*10*,*78*), except that we assume no unobserved selection in our main analysis. In other words, if there is no unobserved selectivity, MTE is reduced to the propensity-score specific effect. To explore the degree to which our propensity score–specific estimates of returns to college are sensitive to unobserved selection, let us now explicitly invoke a generalized Roy model. Consider two potential outcomes, Y_1 and Y_0, a binary indicator D for college completion, and a vector of baseline covariates X. The outcome equations can be written as

$$Y_0 = \mu_0(X) + \epsilon \tag{5}$$

$$Y_1 = \mu_1(X) + \epsilon + \eta \tag{6}$$

where $\mu_0(X) = \mathbb{E}[Y_0|X]$, $\mu_1(X) = \mathbb{E}[Y_1|X]$, the error term ϵ captures all unobserved factors that affect the baseline earnings (Y_0), and the error term η captures all unobserved factors that affect the college premium ($Y_1 - Y_0$). Treatment selection is represented by a latent index model. Let I_D be a latent tendency for college completion, which depends on both observed (X) and unobserved (V) factors

$$I_D = \mu_D(X) - V \tag{7}$$

$$D = \mathbb{I}(I_D > 0) \tag{8}$$

where $\mu_D(X)$ is an unspecified function and V is a latent random variable representing unobserved, individual-specific "resistance to treatment." In the presence of unobserved selection, the latent resistance V may be correlated with ϵ and η, and our estimates of heterogeneous college returns may be biased.

To make the problem analytically tractable, let us further assume that the error terms ϵ, η, and V are jointly Gaussian. Under this assumption, the estimated average treatment effect (ATE) and average treatment effect of the treated (ATT) conditional on the propensity score p are subject to the following biases

$$\text{Bias\_ATE}(p) = -\left[\frac{\rho_{\epsilon V}\sigma_\epsilon + \rho_{\eta V}\sigma_\eta}{p} + \frac{\rho_{\epsilon V}\sigma_\epsilon}{1-p}\right]\phi\left[\Phi^{-1}(p)\right] \tag{9}$$

$$\text{Bias\_ATT}(p) = -\left[\frac{\rho_{\epsilon V}\sigma_\epsilon}{p} + \frac{\rho_{\epsilon V}\sigma_\epsilon}{1-p}\right]\phi\left[\Phi^{-1}(p)\right] \tag{10}$$

where σ_ϵ is the SD of ϵ, σ_η is the SD of η, $\rho_{\epsilon V}$ is the correlation coefficient between ϵ and V, $\rho_{\eta V}$ is the correlation coefficient between η and V, and ϕ and Φ represent the probability density function and the cumulative distribution function of a standard normal distribution, respectively. We see that the bias for ATT(p) depends on two parameters, $\rho_{\epsilon V}$ and σ_ϵ.

In practice, it is difficult to estimate these error variances and correlations without strong instrumental variables (*79*), which are difficult to justify without experimental or quasi-experimental data. To investigate how Bias_ATT(p) changes as a function of the error correlation $\rho_{\epsilon V}$, we first obtain a crude estimate of σ_ϵ using the residual SD from a nonparametric regression of log hourly wage at age 50 on the pretreatment covariates among noncollege graduates (0.65 for men, 0.75 for women). Specifically, we use Bayesian additive regression trees (*80*). This approach serves only as an approximation. When $\rho_{\epsilon V} \neq 0$, our estimates of σ_ϵ are likely biased. Then, for a number of $\rho_{\epsilon V}$ values ranging from −0.2 to 0.2, we calculate the bias-adjusted estimates of the propensity score–specific returns to college for men and women at age 50. Because we are primarily interested in heterogeneity of the treatment effect, our findings will be biased more by a differential violation of unconfoundedness between low- and high-propensity individuals than by a violation of unconfoundedness per se. Because our sensitivity analysis uses a bias formula that implies a larger bias for low- and high-propensity individuals than for individuals in the middle of the propensity score distribution (see Figure S6), it constitutes a particularly conservative test for the robustness of our finding.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abg7641

REFERENCES AND NOTES

1. P. Attewell, D. Lavin, T. Domina, T. Levey, *Passing the Torch: Does Higher Education for the Disadvantaged Pay Off Across the Generations?* (Russell Sage Foundation, 2007).
2. J. E. Brand, Y. Xie, Who benefits most from college? *Am. Sociol. Rev.* **75**, 273–302 (2010).
3. P. Carneiro, J. J. Heckman, E. J. Vytlacil, Estimating marginal returns to education. *Am. Econ. Rev.* **101**, 2754–2781 (2011).
4. S. Dale, A. B. Krueger, "Estimating the return to college selectivity over the career using administrative earnings data" (Technical Report, National Bureau of Economic Research, 2011).
5. S. D. Zimmerman, The returns to college admission for academically marginal students. *J. Labor Econ.* **32**, 711–754 (2014).
6. G. S. Becker, *Human Capital and the Personal Distribution of Income: An Analytical Approach* (Institute of Public Administration, 1964).
7. R. J. Willis, S. Rosen, Education and self-selection. *J. Polit. Econ.* **87**, S7–S36 (1979).
8. M. Hout, Social and economic returns to college education in the United States. *Annu. Rev. Sociol.* **38**, 379–400 (2012).
9. S. B. Dale, A. B. Krueger, Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *Q. J. Econ.* **117**, 1491–1527 (2002).
10. X. Zhou, Y. Xie, Marginal treatment effects from a propensity score perspective. *J. Polit. Econ.* **127**, 3070–3084 (2019).
11. X. Zhou, Y. Xie, Heterogeneous treatment effects in the presence of self-selection: A propensity score perspective. *Sociol. Methodol.* **50**, 350–385 (2020).
12. A. Bernhardt, M. Morris, M. S. Handcock, M. A. Scott, *Divergent Paths: Economic Mobility in the New American Labor Market* (Russell Sage Foundation, 2001).
13. S. Cheng, A life course trajectory framework for understanding the intracohort pattern of wage inequality. *Am. J. Sociol.* **120**, 633–700 (2014).
14. D. Dannefer, Cumulative advantage/disadvantage and the life course: Cross-fertilizing age and social science theory. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **58**, S327–S337 (2003).
15. T. A. DiPrete, G. M. Eirich, Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annu. Rev. Sociol.* **32**, 271–297 (2006).
16. G. S. Becker, Investment in human capital: A theoretical analysis. *J. Polit. Econ.* **70**, 9–49 (1962).
17. H. C. Connolly, P. T. Gottschalk, Differences in wage growth by education level: Do less-educated workers gain less from work experience? IZA DP No. 2331 (2006); https://www.iza.org/publications/dp/2331/differences-in-wage-growth-by-education-level-do-less-educated-workers-gain-less-from-work-experience.
18. J. J. Heckman, L. J. Lochner, P. E. Todd, "Fifty years of mincer earnings regressions" (Technical Report, National Bureau of Economic Research, 2003).
19. C. Sanders, C. Taber, Life-cycle wage growth and heterogeneous human capital. *Annu. Rev. Econ.* **4**, 399–425 (2012).
20. C. Tilly, *Durable Inequality* (University of California Press, 1999).
21. D. Tomaskovic-Devey, D. Avent-Holt, C. Zimmer, S. Harding, The categorical generation of organizational inequality: A comparative test of Tilly's durable inequality. *Res. Soc. Stratif. Mobil.* **27**, 128–142 (2009).

22. S. Haider, G. Solon, Life-cycle variation in the association between current and lifetime earnings. *Am. Econ. Rev.* **96**, 1308–1320 (2006).
23. S. DeLuca, S. Clampet-Lundquist, K. Edin, *Coming of Age in the Other America* (Russell Sage Foundation, 2016).
24. R. Breen, H. van de Werfhorst, M. M. Jaeger, Deciding under doubt: A theory of risk aversion, time discounting preferences, and educational decision-making. *Eur. Sociol. Rev.* **30**, 258–270 (2014).
25. F. Koenig, W. Swanson, C. Harter, Future orientation and social status. *Percept. Mot. Skills* **51**, 927–930 (1980).
26. E. C. Lawrance, Poverty and the rate of time preference: Evidence from panel data. *J. Polit. Econ.* **99**, 54–77 (1991).
27. J.-E. Nurmi, How do adolescents see their future? A review of the development of future orientation and planning. *Dev. Rev.* **11**, 1–59 (1991).
28. A. O'Rand, R. A. Ellis, Social class and social time perspective. *Soc. Forces* **53**, 53–62 (1974).
29. L. C. Guthrie, S. C. Butler, M. M. Ward, Time perspective and socioeconomic status: A link to socioeconomic disparities in health? *Soc. Sci. Med.* **68**, 2145–2151 (2009).
30. D. H. Autor, F. Levy, R. J. Murnane, The skill content of recent technological change: An empirical exploration. *Q. J. Econ.* **118**, 1279–1333 (2003).
31. C. S. Fischer, M. Hout, *Century of Difference: How America Changed in the Last One Hundred Years* (Russell Sage Foundation, 2006).
32. P. Oreopoulos, U. Petronijevic, Making college worth it: A review of the returns to higher education. *Future Child.* **23**, 41–65 (2013).
33. M. Bhuller, M. Mogstad, K. G. Salvanes, Life-cycle earnings, education premiums, and internal rates of return. *J. Labor Econ.* **35**, 993–1030 (2017).
34. J. J. Heckman, L. Lochner, C. Taber, Explaining rising wage inequality: Explanations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Rev. Econ. Dynam.* **1**, 1–58 (1998).
35. R. Topel, Specific capital, mobility, and wages: Wages rise with Job seniority. *J. Polit. Econ.* **99**, 145–176 (1991).
36. K. A. Weeden, Why do some occupations pay more than others? Social closure and earnings inequality in the United States. *Am. J. Sociol.* **108**, 55–101 (2002).
37. R. Chetty, J. N. Friedman, E. Saez, N. Turner, D. Yagan, "Mobility report cards: The role of colleges in intergenerational mobility" (Technical Report, National Bureau of Economic Research, 2017).
38. M. Hout, More universalism, less structural mobility: The American occupational structure in the 1980s. *Am. J. Sociol.* **93**, 1358 (1988).
39. F. Torche, Is a college degree still the great equalizer? Intergenerational mobility across levels of schooling in the United States. *Am. J. Sociol.* **117**, 763–807 (2011).
40. J. E. Brand, J. Xu, B. Koch, P. Geraldo, Uncovering sociological effect heterogeneity using tree-based machine learning. *Sociol. Methodol.* **51**, 189–223 (2021).
41. E. A. Armstrong, L. T. Hamilton, *Paying for the Party* (Harvard Univ. Press, 2013).
42. J. M. Stuber, *Inside the College Gates: How Class and Culture Matter in Higher Education* (Lexington Books, 2011).
43. L. A. Rivera, *Pedigree: How Elite Students Get Elite Jobs* (Princeton Univ. Press, 2016).
44. J. O. Jonsson, D. B. Grusky, M. Di Carlo, R. Pollak, M. C. Brinton, Microclass mobility: Social reproduction in four countries. *Am. J. Sociol.* **114**, 977–1036 (2009).
45. X. Zhou, Equalization or selection? Reassessing the "Meritocratic Power" of a college degree in intergenerational income mobility. *Am. Sociol. Rev.* **84**, 459–485 (2019).
46. S. Cheng, The accumulation of (dis)advantage. *Am. Sociol. Rev.* **81**, 29–56 (2016).
47. P. Cortés, J. Pan, "Children and the remaining gender gaps in the labor market" (Technical Report, National Bureau of Economic Research, 2020).
48. P. England, A. Levine, E. Mishel, Progress toward gender equality in the United States has slowed or stalled. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 6990–6997 (2020).
49. P. England, J. Bearak, M. J. Budig, M. J. Hodges, Do highly paid, highly skilled women experience the largest motherhood penalty? *Am. Sociol. Rev.* **81**, 1161–1189 (2016).
50. J. E. Brand, D. Davis, The impact of college education on fertility: Evidence for heterogeneous effects. *Demography* **48**, 863–887 (2011).
51. K. Musick, J. E. Brand, D. Davis, Variation in the relationship between education and marriage: Marriage market mismatch? *J. Marriage Fam.* **74**, 53 (2012).
52. R. Breen, S. Choi, A. Holm, Heterogeneous causal effects and sample selection bias. *Sociol. Sci.* **2**, 351–369 (2015).
53. G. W. Imbens, D. B. Rubin, *Causal Inference in Statistics, Social, and Biomedical Sciences* (Cambridge Univ. Press, 2015).
54. E. A. Stuart, D. B. Rubin, *Best Practices in Quantitative Methods* (SAGE Publishing, 2008), pp. 155–176.
55. Z. Griliches, Estimating the returns to schooling: Some econometric problems. *Econometrica*, 1–22 (1977).
56. B. Backes, H. J. Holzer, E. D. Velez, Is it worth it? Postsecondary education and labor market outcomes for the disadvantaged. *IZA J. Labor Policy* **4**, 1 (2015).
57. D. J. Henderson, S. W. Polachek, L. Wang, Heterogeneity in schooling rates of return. *Econ. Educ. Rev.* **30**, 1202–1214 (2011).
58. E. R. Eide, M. H. Showalter, Factors affecting the transmission of earnings across generations: A quantile regression approach. *J. Human Resources* **34**, 253–267 (1999).
59. D. Card, Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica* **69**, 1127–1160 (2001).
60. M. J. Budig, P. England, The wage penalty for motherhood. *Am. Sociol. Rev.* **66**, 204 (2001).
61. S. J. Correll, S. Benard, I. Paik, Getting a job: Is there a motherhood penalty? *Am. J. Sociol.* **112**, 1297–1339 (2007).
62. P. Gonalons-Pons, C. R. Schwartz, K. Musick, Changes in couples' earnings following parenthood and trends in family earnings inequality. *Demography* **58**, 1093–1117 (2021).
63. J. J. Heckman, Sample selection bias as a specification error. *Econometrica* **47**, 153–161 (1979).
64. W. H. Sewell, R. M. Hauser, W. C. Wolf, Sex, schooling, and occupational status. *Am. J. Sociol.* **86**, 551–583 (1980).
65. D. Bloome, S. Dyer, X. Zhou, Educational inequality, educational expansion, and intergenerational income persistence in the United States. *Am. Sociol. Rev.* **83**, 1215–1253 (2018).
66. C. Goldin, L. Katz, *The Race Between Education and Technology* (Belknap Press, 2007).
67. C. Manski, D. Wise, *College Choice in America* (Harvard Univ. Press, 1983).
68. S. E. Black, J. T. Denning, J. Rothstein, "Winners and losers? The effect of gaining and losing access to selective colleges on education and labor market outcomes" (Technical Report, National Bureau of Economic Research, 2020).
69. K. Voss, M. Hout, K. George, "College dropout in the United States, 1980–2010: Persistent problems and inequalities" (Technical Report, International Sociological Association Research Committee on Stratification and Mobility (RC28), University of Pennsylvania, 2015).
70. J. E. Brand, J. Xu, B. Koch, P. Geraldo, Uncovering sociological effect heterogeneity using machine learning. arXiv:1909.09138 [stat.OT] (18 September 2019).
71. A. Chandra, "Is the convergence of the racial wage gap illusory?" (Technical Report, National Bureau of Economic Research, 2003).
72. B. Western, B. Pettit, Black-white wage inequality, employment rates, and incarceration. *Am. J. Sociol.* **111**, 553–578 (2005).
73. P. Carneiro, K. T. Hansen, J. J. Heckman, 2001 Lawrence R. Klein lecture estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice*. *Int. Econ. Rev.* **44**, 361–422 (2003).
74. J. J. Heckman, J. E. Humphries, G. Veramendi, The nonmarket benefits of education and ability. *J. Human Capital* **12**, 282 (2018).
75. J. Cawley, K. Conneely, J. Heckman, E. Vytlacil, *Intelligence, Genes, and Success* (Springer, 1997), pp. 179–192.
76. T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, 2009).
77. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models Usinglme4. *J. Stat. Softw.* **67**, 1 (2015).
78. P. Eisenhauer, J. J. Heckman, E. Vytlacil, The generalized Roy model and the cost-benefit analysis of social programs. *J. Polit. Econ.* **123**, 413–443 (2015).
79. X. Zhou, Y. Xie, Propensity score-based methods versus MTE-based methods in causal inference: Identification, estimation, and application. *Sociol. Methods Res.* **45**, 3 (2016).
80. H. A. Chipman, E. I. George, R. E. McCulloch, BART: Bayesian additive regression trees. *Ann. Appl. Stat.* **4**, 266 (2010).

Acknowledgments: We thank the *Science Advances* reviewers for their constructive comments. Earlier versions of this paper were presented at the Plenary Session of the 2019 Summer Meeting of the International Sociological Association's Research Committee on Social Stratification and Mobility (RC28) in Princeton, New Jersey and the 2018 Annual Meeting of the Population Association of America in Denver, Colorado. **Funding:** J.E.B. benefited from facilities and resources provided by the California Center for Population Research at UCLA (CCPR), which receives core support (P2C-HD041022) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). Y.X. and J.E.B. received financial support for this research from NIH grant R01 HD07460301A1. **Author contributions:** All authors contributed equally to the project. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Replication materials are available in Open Science Framework: https://osf.io/vkt8y/.

Submitted 31 January 2021
Accepted 27 October 2021
Published 15 December 2021
10.1126/sciadv.abg7641