NBER WORKING PAPER SERIES

HUMAN CAPITAL AND BLACK-WHITE EARNINGS GAPS, 1966-2017

Owen Thompson

Working Paper 28586
http://www.nber.org/papers/w28586

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021

I thank the Upjohn Institute for Employment Research for financial support and seminar participants at Amherst College, Notre Dame and SUNY-Albany for helpful comments. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Human Capital and Black-White Earnings Gaps, 1966-2017
Owen Thompson
NBER Working Paper No. 28586
March 2021
JEL No. I24,I26,J71

ABSTRACT

This paper estimates the contribution of human capital, measured using both educational attainment and test performance, to the Black-white earnings gap in three separate samples of men spanning 1966 through 2017. There are three main findings. First, the magnitude of reductions in the Black-white earnings gap that occur after controlling for human capital have become much larger over time, suggesting a growing contribution of human capital to Black-white earnings disparities. Second, these increases are almost entirely due to growth in the returns to human capital, rather than changing racial gaps in the human capital traits themselves. Finally, growth in the explanatory power of human capital has been primarily due to increases in the association between human capital and the likelihood of non-work, with no clear increases in the extent to which human capital explains Black-white differences in hourly wages or other intensive margins. These findings highlight how apparently race-neutral structural developments in the US labor market, such as increasing skill prices and falling labor force participation rates among less skilled men, have had large impacts on the dynamics of racial inequality.

Owen Thompson
Department of Economics
Williams College
24 Hopkins Hall Drive
Williamstown, MA 01267
and NBER
ot3@williams.edu

A data appendix is available at http://www.nber.org/data-appendix/w28586

Introduction

Large disparities in the earnings of African Americans and whites have existed for as long as data on labor market outcomes has been collected, and remain one of the defining social and economic issues of the United States (Mydral 1944; Altonji & Blank 1999; Lang & Lehmann 2012).

Among the many explanations for these disparities that have been advanced by researchers and policy actors, a fundamental distinction can be drawn between explanations that emphasize differences in the human capital levels of Blacks and whites versus explanations that emphasize differential treatment in the labor market among similarly qualified individuals of different races. This distinction is important in part because it is viewed as having direct policy implications: If earnings disparities are primarily traceable to human capital disparities, it is assumed that policies which promote human capital acquisition among minorities will reduce earnings gaps. In contrast, if human capital is not the primary driver of Black-white earnings gaps, then policies that expand or more vigorously enforce anti-discrimination laws may be more effective.

This paper provides a broad assessment of human capital's contribution to racial earnings disparities from the period following the 1964 Civil Rights Act, when meaningful protections from overt labor market discrimination were first established, through the present. Using three nationally representative samples of men with information on both formal schooling and standardized test performance, I find that differences in the human capital levels of Black and white men "explain" a large and increasing share of racial earnings disparities: Controlling for human capital reduces Black-white differences in total earnings by approximately 1 log-point (10%) in data from the late 1960s and the 1970s, by approximately 27 log points (15%) in data from the 1980s and 1990s, and by approximately 55 log points (30%) in data from the 2000s and 2010s.

While the overall contribution of human capital to racial earnings disparities has unambiguously increased in recent decades, the precise reasons for these increases are nuanced and perhaps surprising. In particular, I show that two features of the increasing explanatory power of human capital are especially important for understanding the relationship between human capital and Black-white earnings gaps.

First, the increased explanatory power of educational attainment and standardized test performance is almost wholly due to growth in the earnings premiums associated with these traits. Increases in skill premia starting in the late 20th Century are well established and are the topic of an extensive and active literature (see Katz & Murphy 1992; Autor, Katz & Kearney 2008; and Goldin & Katz 2009, 2020 among many others). The current paper's findings highlight that in the presence of these rising skill premia, a large overall contribution of human capital to the earnings gap does not necessarily imply that policies which increase the relative test scores or educational attainment of African Americans will also reduce the Black-white earnings gap. Indeed the large increases in human capital's explanatory power found here occurred in a context of constant or declining Black-white gaps in education and test scores.

Second, the increasing importance of human capital is driven almost entirely by the extensive margin between work and non-work. At intensive margins such as hourly wages and hours worked among employed men, the importance of human capital to Black-white disparities has actually fallen significantly between the 1980s and the present. The growing importance of non-work is closely related to the widely studied phenomena of declining labor force participation rates among less-skilled men over the past several decades, as well as to increasingly punitive criminal justice policies that have had a highly disproportionate impact on Black men (see Autor & Duggan 2003; Neal & Rick 2014; Krueger 2017; and Aguiar et al. 2017).

The paper's findings contribute to two distinct strands of the extensive existing literature on US racial inequality.

First, and most directly relevant, is a set of studies that estimate the contribution of human capital to racial gaps in labor market outcomes. A canonical paper in this area is Neal & Johnson (1996), who report that controlling for standardized test scores alone eliminates a large share of the Black-white gap in hourly wages, and other studies using similar methodological approaches (but sometimes reaching different conclusions) include Lang & Manove (2011); Ritter & Taylor (2011); Fryer (2011); Gelbach (2016); Carruthers & Wanamaker (2017) and Luo (2020). One contribution of the current paper is to build on this literature by providing an analysis with a similar methodology as Neal & Johnson (1996) and these subsequent studies, but with samples that span a wider range of years and with a focus on additional outcome measures, especially the extensive margin.

Second, and perhaps more importantly, I contribute to a recent strand of the literature that emphasizes how structural aspects of the labor market can impact racial earnings inequality. For instance recent work by Bayer & Charles (2018) evaluates how changes in the overall earnings distribution contributed to differences in the earnings of Black and white men between 1940 and the present, with a key finding being that racial earnings inequality at the median has been predominantly determined by the earnings trajectories of *all* men in the lower percentiles of the earnings distribution, where Black men are disproportionately concentrated. Another important study in this area is Derenoncourt & Montialoux (2018), who show that a large share of the racial convergence in hourly wages that occurred in the late 1960s and early 1970s can be attributed to minimum wage coverage expansions included in the 1966 Fair Labor Standard Act, which were technically race neutral but that in practice disproportionately benefited African American workers. Additionally, Chetty et al. (2020) estimate how intergenerational income transmission varies across racial groups, and study how these intergenerational dynamics affect the long term structural determination of racial inequality.[1]

These recent studies are very informative about the structural evolution of US racial inequality, but generally do not directly evaluate the role of individual skill differences. The current paper builds on this strand of the literature by incorporating more direct and comprehensive human capital measures over a wider range of cohorts, which in particular allows me to better evaluate how racial inequality has been affected by long

[1] An older literature studies the relationship between the overall US wage structure and racial inequality (Card & Lemieux 1994, 1996; Maloney 1994; Reardon 1997), but does not go beyond the 1980s, does not consider the extensive margin, and does not directly incorporate skill measures other than educational attainment.

term trends like rising skill premia and falling male labor force participation rates.

The remainder of the paper proceeds in four sections. Section 1 describes the data and measures; Section 2 reports the baseline results and establishes their robustness and reliability; Section 3 reports the results of several extensions; and Section 4 concludes.

1 Data and Measures

I use three separate nationally representative longitudinal samples, all of which contain detailed information on earnings, educational attainment and standardized test scores, in addition to basic demographic information. This section describes these samples and the key variables used in the analysis.

1.1 Samples

The first data set is the National Longitudinal Survey "Original Cohort" of Young Men (NLS-OC). The NLS-OC tracked a nationally representative sample of 5,225 men who were born between 1941 and 1952, interviewing respondents 12 times between 1966 and 1981.

The second data set is the 1979 National Longitudinal Survey of Youth (NLSY-79), which includes 6,403 male respondents who were born between 1957 and 1964 and have been surveyed on an annual or bi-annual basis from 1979 through the present.

The third and final data set is the 1997 National Longitudinal Survey of Youth (NLSY-97), whose 4,599 male respondents were born between 1980 and 1984 and have been interviewed annually or bi-annually from 1997 through the present, with the most recent available wave occurring in 2017. Participants in the NLSY-97 were ages 32-37 in the most recent available wave, and therefore only recently reached an age range where adult labor market outcomes can be meaningfully observed.

I restrict all three samples to men ages 21-37, and exclude years in which respondents reported being enrolled in school. The lower bound on age and the exclusion of current students are imposed to help ensure that respondents have substantively entered the labor force at the time of observation, while the upper age bound of 37 is the age of the oldest NLSY-97 respondents in the most recent wave. Previous research has found that earnings observed in this age range are reasonably representative of lifetime earnings (Black & Devereux 2010; Chetty et al. 2014; Mazumder 2018), and below I show that the paper's main findings are not sensitive to alternative age ranges or to including men currently enrolled in school.[2]

[2] All three surveys contain over-samples of Blacks and provide sampling weights designed to make the samples nationally representative. I apply these weights in the main analyses, and demonstrate robustness to not applying sampling weights in Section 2.2. The NLSY-79 originally contained over-samples of military personnel and low-income whites, in addition to racial minorities, but these were discontinued in 1984 and 1990 for budgetary reasons and are excluded here. I also restrict all three samples to non-Hispanic Blacks and whites, and exclude other racial and ethnic minorities.

After applying these age restrictions, the labor market outcomes of men from the NLS-OC are observed primarily in the 1970s, the labor market outcomes men from the NLSY-79 are observed primarily in mid-1980s through the late-1990s, and the labor market outcomes men from the NLSY-97 are observed primarily in the mid-2000s through the present. The three data sets can therefore jointly provide evidence on the nature of Black-white earnings dynamics throughout the post Civil Rights period.

Of the three surveys, the data from the NLS-OC is unambiguously of the lowest quality. This is true both with respect to certain aspects of the overall sample construction, and for the measurement of some of the key variables, most importantly standardized test scores. The relevant data quality issues are discussed in detail in Section 1.2 below, and estimates that make various adjustments to the NLS-OC to test whether data limitations are likely to alter the paper's main conclusions are reported in Section 2.3 and in the online appendix. Overall these exercises indicate that while the NLS-OC has unambiguous shortcomings, the survey still allows for estimates that are both internally credible and generally comparable to the estimates from the other two surveys.

1.2 Measures

Earnings Measures

My preferred earnings measure is the total of all income from wages, salaries, farm and business income in the previous calendar year, which was reported in all waves of all three surveys, with men who have zero earnings included. This measure is maximally holistic since it incorporates both wage levels and the probability and intensity of labor force participation.

Annual earnings are inflated to 2017 dollars then transformed using the inverse hyperbolic sine function, which allows for an interpretation of coefficients that is similar to the interpretation of log earnings in most applications, but preserves individuals with zero earnings (Burbidge, Magee & Robb 1988; Bellemare & Wichman 2020). For expositional convenience I refer to the inverse hyperbolic sine of earnings as "log" earnings. In Section 2.2 I show that the results are similar if I measure earnings as ln(earnings+1), and in Section 3.1 I report specifications that use a binary measure of positive earnings and a respondent's hourly wages as the outcome measure.[3]

To validate the earnings measures from the relatively small surveys used here, as well as to characterize the general trends in earning disparities over the study period, Figure 1 plots the annual Black-white gap in log total earnings in the NLS-OC, NLSY-79 and NLSY-97, as well as comparably constructed estimates from the Decennial Censuses and ACS. Reassuringly, Figure 1 shows that the levels and trends of the Black-white earnings gap in the longitudinal surveys used here are very similar to those observed in the much

[3] Hourly wages were reported directly by respondents or inferred by NLS staff based on respondent's total compensation, time unit of pay, and total hours, and correspond to their primary (or "CPS") job. Hourly wage values below \$3 and above \$100 are trimmed, and the wage variable is transformed using the natural log function, rather than the inverse hyperbolic sine function.

larger Census and ACS samples. In both data sources, racial disparities in total earnings are large and are increasing over the study period, going from approximately 50 log points in the early 1970s to approximately 200 log points in the 2010s. The large magnitudes of these gaps and their growth over time may seem surprising, but highlight the fact that including zero-earners qualitatively changes the nature of Black-white earnings gap trends.

Human Capital Measures

The educational attainment of each respondent is measured using their highest grade completed at the time of each survey wave, and in some specifications this continuous measure is re-coded into categorical variable indicating education less than high school, high school completion, some college, and college competion or beyond.[4]

A unique feature of all three surveys is that in addition to formal schooling they contain credible standardized test score measures for a large number of respondents. Relative to measuring human capital only with formal schooling, incorporating standardized test performance allows me to partially account for otherwise difficult to observe characteristics like cognitive ability, school quality, and family background.

For the NLSY-79 and NLSY-97 samples, I use scores on the Armed Forces Qualification Test (AFQT), which was administered directly by the survey administrators to 93.9% of NLSY-79 participants and to 79.3% of NLSY-97 participants, and which has been widely used and validated in the economic and psychological literatures. I use AFQT scores that were adjusted for age at the time of testing by survey administrators, and express these scores in standardized units (z-scores).[5]

No standardized tests were administered to NLS-OC participants directly by the NLS survey enumerators. However, the high schools attended by NLS-OC participants were surveyed in 1968, and among other items these school surveys collected the results of any available standardized tests taken by survey participants. Of the 4,007 NLS-OC participants for whom standardized test scores were sought, their secondary schools were able to provide scores for 3,375 young men, a response rate of 84%.[6] Data on over 30 different standardized tests was collected, with the most common specific tests being the Otis/Beta/Gamma Test (848 respondents), the California Test of Mental Maturity (625 respondents), the Preliminary Scholastic Aptitude Test (223 respondents) and the Henmon-Nelson Test (216 respondents). Working with NLS administrators, Herriott & Kohen (1973) then collected information on the means and standard deviations of each test from the test publishers, and used these moments to convert the raw scores from the school survey onto a common scale.[7]

[4] The NLS-OC did not collect data on degree completion, and so to maximize comparability across surveys I base the education variables for all three surveys on continuous highest grade completed measures.

[5] Altonji, Bharadwaj & Lange (2012) develop a detailed methodology for making AFQT scores comparable across the NLSY-79 and NLSY-97 which accounts for pencil-paper versus computer based testing formats across the two surveys and other factors in addition to age-at-testing adjustments. Because the current study is primarily concerned with the extent to which test scores can explain racial earnings differences *within* each sample, while Altonji, Bharadwaj & Lange (2012) are primarily concerned with estimating how the characteristics of youth changed across the cohorts represented in the two surveys, I do not apply a comparable adjustment and instead use the age-adjusted AFQT scores provided directly by the NLSY survey administrators.

[6] The full NLS-OC Young Men's sample contained 5,225 respondents, but test scores were only sought for 4,007 men, with a large majority of these exclusions being young men who had not not yet completed 9th grade at the time of the school survey.

[7] It is not entirely clear what populations were used by the various test publishers when calculating these means and standard

An obvious potential issue is that the test scores available in the NLS-OC are measured with substantial error or are sufficiently different from those available in the later surveys that comparisons across the data sets are uninformative or misleading. One relatively direct assessment of the comparability of the test score measures across the three surveys is made possible by the fact that in 1980 the NLSY-79 also conducted a survey of the secondary institutions attended by its participants, and collected scores on various standardized tests scores for a subset of NLSY-79 participants. This allows for a direct comparison of AFQT scores and school-survey derived test scores within a common sample. In the online appendix I show that in earnings equations, the coefficients on the two types of test scores are very similar, and in Section 2.3 below I show that if the test score coefficients in the NLS-OC are scaled by a first-stage regression of school survey derived test scores onto AFQT scores estimated with NLSY-79 data, the study's main findings are not qualitatively changed.[8]

Another potentially important data quality issue arises from the fact that the NLS-OC collected test scores from the high schools attended by respondents, so that data on test performance is uniformly missing for respondents whose educational attainment did not advance beyond primary school, and is more likely to be missing for individuals who completed fewer years of high school. Because Black respondents were less likely to enter and complete high school, excluding respondents without test scores decreases the measured Black-white gap in years of completed education. In Section 2.3 I report results that adjust for the effects of truncating the NLS-OC sample in this way, and also report results from the NLSY-79 and NLSY-97 when those samples are modified to have similar truncation as the NLS-OC sample, and both exercises suggest that the study's main findings are not qualitatively affected by this feature of the NLS-OC sample.

2 Human Capital and Black-White Differences in Total Earnings

2.1 Main Findings

My basic empirical approach simply compares the magnitude of the unconditional Black-white earnings gap with the magnitude of the earnings gap conditional on human capital characteristics. I specifically estimate regressions of the following form separately in the NLS-OC, NLSY-79 and NLSY-97 samples:

deviations. In the online appendix I show that after standardization, tests which seem likely to have been taken by a positively selected subset of NLS-OC respondents (e.g. the National Merit Scholarship Qualifying Test) are associated with higher standardized test scores. I also report the results of models that include test-type fixed effects, and the returns to standardized test scores in the NLS-OC are virtually unchanged in these specifications. These results suggests that the norming samples used by test publishers to generate means and standard deviations were not simply the individuals naturally taking each particular test, although how representative they are of the general population is not clear.

[8] In addition to these assessments, Herriott & Kohen (1973) discuss the validity of pooling disparate test measures in the NLS-OC at length, including estimating whether the associations between test performance and parental education and occupation vary across different tests. The authors find that socioeconomic background has similar effects on test performance across the available test types, and conclude that "we see little reason for social scientists... to be reluctant to pool data from different commonly used tests of mental ability." I also note that several prior studies have compared the effects of the test score measures from the NLS-OC with those in the later NLSY surveys, including Cunha & Heckman (2016) and Bacolod & Hotz (2006).

$$Earnings_i = \alpha + \beta_1 Black_i + X_i\delta + \varepsilon_i$$

where $Earnings_i$ is the log labor market earnings of individual i; $Black_i$ is an indicator for whether individual i is Black rather than white; and X_i is a vector of individual level controls. In the baseline estimates of Equation 1, the control vector is empty, while in later specifications it contains human capital controls. I first present estimates using total earnings with zeros included, and later report estimates with alternative dependent variables that focus on specific earnings margins.

The baseline unconditional estimates for total earnings are reported for each of the three data sets in Columns 1, 3 and 5 of Table 1. These results indicate that the unadjusted Black-white earnings differential was 96 log points in the NLS-OC sample, grew to 168 log points in the NLSY-79 sample, and grew further to 197 log points in the NLSY-97 sample. These large and growing earnings gaps are in line with the patterns shown in Figure 1.

Next I add controls for standardized test scores and educational attainment to the X_i vector. Adding these covariates refines the earnings comparisons to Black and white men with similar levels of observable human capital, and therefore any attenuation in the Black indicator provides a descriptive estimate for how much of the Black-white earnings gap is attributable to human capital differences.

The conditional estimates are reported in Columns 2, 4 and 6 of Table 1, and the level and percent changes in the Black coefficient after controlling for human capital are reported in the bottom two rows of the table. Unsurprisingly, controlling for human capital reduces the estimated Black-white earnings gap in all three data sets. More notably, the magnitude of the declines in β_1 that occur after conditioning on education and test scores grows over time. In the NLS-OC, controlling for human capital characteristics decreases the estimated racial earnings gap from .960 to .866, a decrease of 9 log points or 9.8%. In the NLSY-79, adding the same controls decreases the estimated gap from 1.681 to 1.407, a decrease of 27 log points or 16.3%. Finally, in the NLSY-97 the controls for education and standardized test scores reduce the Black-white earnings differential from 1.974 to 1.420, a decrease of 55 log points or 28.1%.

These increases in the explanatory power of human capital constitute one of the current paper's main findings. However, these results do not readily differentiate the impact of educational attainment versus standardized test scores. This distinction is non-trivial, since some influential prior studies, most notably Neal & Johnson (1996), have found that performance on standardized tests rather than educational attainment is the dominant factor in explaining racial wage differences.[9] One common practice for attempting to determine which specific control variables are driving attenuation in an independent variable of interest is to add the covariates sequentially, for instance by first controlling for test scores, and then additionally controlling for education. But the findings of such an exercise will be dependent on the order in which the covariates are added, and this ordering choice is arbitrary in most applications including the current one.[10]

[9] Other studies finding that test performance strongly outweighs formal schooling include Fryer (2011) and Luo (2019).

[10] Indeed Lang & Manove (2011) show that controlling for education in a model that already conditions on test scores actually

As an alternative, I implement the decomposition method proposed by Gelbach (2016), which allows for the reduction in a coefficient of interest that is attributable to each specific control variable to be estimated in a manner that is invariant to the order in which the covariates are added.

In particular Gelbach (2016) shows that the contribution of a particular covariate to the reduction in a coefficient of interest will be equal to the product of two easily estimable parameters. First is the covariate's coefficient in the model with the full control vector, which in the current application are the coefficients on test scores and education that were reported in the even numbered columns Table 1. These parameters provide an estimate of the covariate's association with the dependent variable, conditional on all other covariates. Second is the coefficient on the independent variable of interest in an auxiliary regression of the covariate onto the independent variable of interest. In the current application, these are regressions of test scores and education onto a Black indicator, and therefore estimate the Black-white gap in these characteristics.[11] The contribution of each covariate is simply the product of these two parameters, and the Gelbach decomposition therefore formalizes the intuition that the extent to which a human capital measure can "explain" Black-white earnings differences depends jointly on the measure's conditional association with earnings, *and* on how strongly the human capital measure differs by race. Because the *conditional* effect of each covariate is used in this decomposition, the results do not depend on the arbitrary choice of which covariates are added first.

The first key set of parameters for the Gelbach (2013) decomposition, the conditional returns to test scores and educational attainment, were already reported in Columns 2, 4 and 6 of Table 1. These estimates indicate that the returns to both standardized test scores and formal schooling increased substantially across the three data sets, consistent with a large existing literature on increases in skill prices in recent decades. Specifically, within the NLS-OC an additional year of schooling was conditionally associated with a 7.2 log point increase in total earnings, while this association grew to 15.7 log points in the NLSY-79 and to 19.1 log points in the NLSY-97. Similarly, a standard deviation increase in standardized test performance was conditionally associated with a 2.1 log point increase in earnings in the NLS-OC sample, while this association grew to 13.7 log points in the NLSY-79 and to 41 log points in the NLSY-97.

The second set of key parameters for the decomposition are the racial gaps in each human capital characteristic, which are estimated with supplemental regressions of each characteristic onto a Black indicator. Results of these regressions are reported in Table 2, and in contrast to the rapidly increasing returns to human capital, shows that Black-white differences in human capital were relatively stable across the three surveys. Specifically Columns 1-3 of Table 2 show that the racial gap in years of education fell from 1.01 to .83 years between the NLS-OC and NLSY-79, then grew to 1.12 years in the NLSY-97. Likewise, Columns 4-6 of Table 2 estimate that the Black-white gap in standardized test scores was 1.02 standard deviations in

increases the estimated Black-white wage gap. This occurs because conditional on test scores, Black men actually obtain substantially more education than white men. See Lang & Manove (2011) and Gelbach (2016) for detailed discussions.

[11] Gelbach (2016) shows that the estimated contribution of each covariate calculated in this fashion will sum to the total reduction in the coefficient of interest as an identify, and derives standard error formulas.

the NLS-OC, was virtually unchanged at 1.03 standard deviations in the NLSY-79, then closed moderately to .83 standard deviations in the NLSY-97.

These Black-white disparities in human capital are generally consistent with existing estimates from more authoritative data sources. For instance, while there is not standardized national test score gap data spanning back to the period covered by the NLS-OC, the test score gap among 17 year-olds taking the National Assessment of Educational Progress mathematics test was approximately 1.18 in 1978 and then fell to .96 in 1996.[12] Similarly, in the online appendix I report Black-white gaps in years of educational attainment in the Decennial Census and find a gap of .95 years for Census respondents from the same cohorts and the NLSY-79 sample, and a gap of 1.01 years for Census respondents from the same cohorts as the NLSY-97 sample.

The one estimate from Table 2 that does not closely adhere to patterns in other data sources is the 1.01 year gap in educational attainment in the NLS-OC, which is substantially smaller than in other data sets and likely underestimates the true Black-white education gap for these cohorts. For instance in the online appendix I find a Black-white education gap of 1.58 years for Census respondents from the same cohorts as the NLS-OC sample, and Chay, Guryan & Mazumder (2014) and by Bayer & Charles (2018) report similarly large education gaps from this period. As noted in Section 1 this is likely due to less educated NLS-OC respondents being less likely to have valid test score data and therefore being excluded from the estimation sample. For consistency I first report baseline decomposition results that use the 1.01 year education gap in my actual working NLS-OC sample, and then in Section 2.3 I report adjusted results that use Black-white education gaps in line with consensus estimates as well as results that modify the NLSY-79 and NLSY-97 samples to be more comparable to the NLS-OC sample with respect to educational composition.

Table 3 combines the estimated returns to human capital with the estimated size of human capital gaps and reports full decomposition results for each survey. Column 1 of Table 3 shows that in the NLS-OC, the 9 log point overall reduction in the earnings gap that occurred after controlling for human capital was primarily attributable to education (7 log points) rather than standardized test scores (2 log points). Column 2 shows that in the NLSY-79, where there was a much larger total reduction of 27 log points, the contribution of education and test scores were almost identical, at 13 log points and 14 log points, respectively. Finally, Column 3 shows that the total explanatory power of human capital grew even further to .55 log points in the NLSY-97, and that these more recent increases were driven by increasing importance of both education and test scores, but especially by test scores. Specifically the estimated reduction in the total earning gap due to educational attainment grew to 21 log points while the estimated reduction due to test scores grew to 34 log points.

On balance these decomposition results indicate that while test scores had somewhat more explanatory power than educational attainment in the most recent cohorts, prior to the 2000s formal schooling was as

[12] Mathematics scores retrieved from from nces.ed.gov/nationsreportcard/data/ on 12/19/2020.

important or more important than test performance. This contrasts with some earlier studies that assigned a dominant role to test scores.

2.2 Robustness

There are various reasonable alternatives to the modeling choices made in the baseline specifications, and Figure 2 demonstrates the robustness of the key patterns to a large number of these alternatives.

I specifically demonstrate robustness to four aspects of model selection. First is sample restrictions, where in the baseline specifications I restricted the sample to men age 21-37 and excluded current students, but reasonable alternatives include using a minimum age of 25 to focus on an age range more representative of lifetime earnings, using a minimum age of 32, since the youngest NLSY-97 respondents were 32 at the time of the most recent survey, and expanding the sample to include current students. Second is the functional form, where for the human capital measures the linear functional forms could reasonably be replaced by sets of dummies for total years of education completed and standardized test score quartile, or where total earnings could be transferred as the natural log of total earnings plus one, rather than with the inverse hyperbolic sine function, and still retain zero earners. Third is the inclusion of baseline covariates, where the specifications above did not include any independent variables beyond race indicators and human capital measures, but the estimates may be more precise or stable when applying a vector of baseline controls, with one reasonable choice being age indicators, a south indicator, and an urban residence indicator. Fourth is the application of sampling weights, where especially in the NLS-OC the quality of these weights is unclear, and a transparent alternative is to estimate unweighted models.

Taking all possible permutations of these alternative modeling choices generates 128 possible specifications, and Figure 2 graphically reports the results of these 128 alternative models.

Panel A of Figure 2 shows the reduction in the Black-white earnings gap that occurs after controlling for human capital in each of the three data sets. The reductions reported in Table 1 are shown with a bold line, while the 127 alternatives are shown with light gray lines. The figure shows that the magnitude of the reduction from controlling for human capital grows across the surveys in all of the alternative specifications, and also that the baseline results reported in Table 1 were not outliers relative to the parameters generated by the full set of alternative modeling choices.

Panels B and C of Figure 2 conduct a similar exercise for the other key pattern from Table 1, the increasing returns to human capital across the three data sets. Specifically Panel B plots the coefficients on the educational attainment variable while Panel C plots the coefficients on the standardized test score variable.[13] Panel B shows that the returns to formal schooling were uniformly increasing across alternative specifications, and that the coefficients on educational attainment in the baseline results from Table 1 were on the high

[13] Panels B and C necessarily exclude the alternative specifications with non-linear measures of education and test performance, so that there are 64 possible models rather than 128.

end of the range of estimates generated in the NLS-OC, but were otherwise quite typical. The results for standardized test scores that are shown in Panel C and are somewhat more variable than those for educational attainment, and in eleven of the 64 models actually show a small decline in the returns to test scores between the NLSY-79 and the NLSY-97.[14] However, the overall pattern is clearly one of increasing returns to test performance, and the baseline estimates from Table 1 are again not outliers compared to the full set of alternative specifications.

2.3 Data Quality in NLS-OC

As discussed in Section 1, there are two important data quality issues with the NLS-OC sample: Potential mismeasurement of standardized test performance, and the disproportionate exclusion of NLS-OC respondents with less education, which occurs mechanically because test score data was collected from respondent's high schools. In Tables 4 and 5 I report the results of two exercises that investigate whether these data quality issues are meaningfully influencing the key findings from Tables 1-3.

First, in Table 4 I report the results of an exercise that attempts to make direct adjustments to the relevant parameters in the NLSY-OC that may be affected by data quality.

Specifically, in Column 1 of Table 4 I report a coefficient for test scores in the NLS-OC that adjusts for measurement error. This adjustment is made by exploiting the fact that the NLSY-79 also conducted a survey of the secondary institutions attended by its respondents, and collected scores on various standardized tests in a manner similar to the test score collection approach used in the NLS-OC. Since the NLSY-79 also administered the AFQT directly to respondents, I am able to estimate a first-stage regression of AFQT scores onto test scores collected from high schools, and this first-stage estimate is equal to .731. The test score coefficient in Column 1 of Table 4 simply scales the baseline NLS-OC test score coefficient (.021) by the first-stage estimate (.731), resulting in an adjusted test score coefficient of .029. For reference, Column 1 of Table 4 also reproduces the coefficient on educational attainment in the NLS-OC from Table 1.

In Column 2 of Table 4 I report the Black-white gap in years of education from the NLS-OC when respondents without valid test score observations are included, and the estimated gap in this sample is 1.69 years. This is substantially larger than the 1.01 year gap that was reported in Table 2 for the baseline NLS-OC sample that excluded respondents with missing test score data, and is more similar to the 1.58 year gap observed among Decennial Census respondents from the 1941-1952 birth cohorts, which is reported in the online appendix. Again for reference, Column 2 reproduces the estimated Black-white gap in test scores in the NLS-OC from Table 2.

Column 3 simply multiplies the human capital coefficients from Column 1 by the human capital gaps from Column 2 to produce adjusted estimates of the portion of the Black-white earnings gap attributable to

[14] There are eleven specifications with declining returns to test performance, and nine of these include current students in the sample, and so maybe attributable to men with higher test scores staying in school further into the life cycle in the NLSY-97 than in the NLSY-79.

each human capital characteristic. When using the adjusted figures from Columns 1 and 2, the estimated contribution of educational attainment to the Black-white earnings gap in the NLS-OC is -.122, while the estimated contribution of test scores is -.030. These estimates are moderately larger than the analogous baseline estimates of -.07 and -.02 from Table 3, and the combined contribution of both human capital characteristics when using the adjusted estimates (-.152) is also non-trivially larger than the baseline total (-.094), which indicates that the lower data quality of the NLS-OC likely does have some impact on the estimates.

However, even using the adjusted NLS-OC estimates from Table 4, the total explanatory power of human capital in the NLS-OC (15 log points) is still much smaller than the analogous estimate for the NLSY-79 (27 log points) and for the NLSY-97 (55 log points). The fundamental reason that the qualitative patterns across samples are unchanged, even though the estimated contribution of human capital in the NLS-OC increases in the adjusted estimates, is that the growth in the returns to education and test scores across the three surveys are *very* strong, and the magnitudes of these skill price increases prevents qualitative changes in the conclusions even in the presence of non-trivial bias in the relevant parameter estimates.

A second approach to evaluating whether the NLSY-OC sample truncation in particular may be affecting the main findings is to modify the NLSY-79 and NLSY-97 samples so that they have educational compositions similar to the NLS-OC.

As noted, the test score collection process in the NLS-OC mechanically led less educated respondents to be more likely to have missing test score data. More specifically, in the NLS-OC test score data is missing for all respondents with 8 or fewer years of education, 60% of respondents with 9 years of education, 40% of respondents with 10 years of education, and 33% of respondents with 11 years of education. In Table 5, I report estimates from specifications that are identical to those in Table 1, but that use NLSY-79 and NLSY-97 samples where I have randomly selected and excluded the same share of respondents within these education groups.[15] Intuitively, while I am not able to recover test scores for many of the less educated NLS-OC respondents, I am able to throw out the test scores of similar respondents from the NLSY-79 and NLSY-97, which lowers data quality in the more recent surveys but also makes comparisons across the three surveys more credible.

The results in Table 5 are very similar to the main findings from Table 1. Since the NLS-OC is not modified, we again observe a reduction in the Black-white earnings gap of 9 log points (9.8%). Within the modified NLSY-79 sample, this reduction still increases to 24 log points (15.1%), and within the modified NLSY-97 sample this reduction still increases to 48 log points (29.8%). There also continues to be strongly increasing skill prices across the three surveys.[16] In conjunction with the results in Table 4, these patterns strongly

[15] To mimic the data structure of the NLS-OC respondents are randomly selected for exclusion at the individual level, rather than the person-year level. A simpler but less precise alternative approach is to simply exclude all respondents with 9 or fewer years of education from all three surveys, and such an approach produces very similar results to those in Table 5.

[16] The magnitudes of the Black-white earnings gap in the NLSY-97, both unconditional and conditional, are smaller when many less educated workers are removed from the sample in Table 5, but the change in the Black-white gap after controlling for human capital is very similar.

suggests that while there are meaningful data quality shortcomings in the NLS-OC sample, these issues are unlikely to be a primary explanation for the main findings.

3 Extensions

3.1 Earnings Differences on the Extensive and Intensive Margin

Because the estimates in Table 1 retained zero-earners, they encompass both an extensive margin between work and non-work, as well as an intensive margin of total earnings conditional on working, which in turn depends on hourly wages and total hours worked. Several previous studies have emphasized the importance of accounting for racial differences in non-participation specifically when evaluating racial earnings differences (Heckman, Lyons & Todd 2000; Chandra 2000; Juhn 2003), and the relative importance of these different margins is clearly essential for fully characterizing changes in the contribution of human capital to racial earnings differentials. Table 6 reports additional specifications that help evaluate the role of different earnings margin.

Panel A of Table 6 reports a set of conditional and unconditional specifications similar to those in Table 1, but that use an indicator of having positive earnings as the dependent variable and therefore directly assess the extensive margin between work and non-work. The basic patterns closely mirror those in Table 1 for total earnings, with rapidly growing unconditional Black-white differences in non-work across the three samples, and with human capital "explaining" a progressively larger share of these gaps. Specifically in the NLS-OC there is a relative modest 5 percentage point unconditional Black-white gap in having positive earnings, which is unchanged to two decimal places after conditioning on education and test scores. In the NLSY-79, the initial gap in positive earnings is 10.9 percentage points and this gap falls to 9.9 percentage points with the human capital controls included, a reduction of 1 percentage point or 9.3%. Finally, in the NLSY-97 the unconditional gap in positive earnings grows further to 14.9 percentage points and conditioning on education and test scores reduces this differential to 10.9 percentage points, a decline of 4 percentage points or 26.5%.

Also similar to the total earnings estimates in Table 1, the increasing importance of human capital for explaining Black-white gaps in non-zero earnings primarily reflects increases in the returns to human capital. Specifically the estimates in Panel A indicate that an additional year of schooling increases the probability of non-zero earnings by a statistically insignificant .1 percentage points in the NLS-OC, but this association increases to statistically significant coefficients of .7 percentage points in the NLSY-79 and to 1.2 percentage points in the NLSY-97. Similarly, a standard deviation increase in test scores is virtually uncorrelated with the probability of having non-zero earnings in the NLS-OC, but with a .4 percentage point increase in the probability of positive earnings in NLSY-79, and with a 3.2 percentage point increase in the NLSY-97.

Panel B of Table 6 reports the results from a similar set of specifications that use hourly wages as the outcome variable. Hourly wages are a key component of the intensive margin, and are also the most widely used outcome measure in the literature on racial disparities in economic outcomes, so that the results in Panel B are useful for comparisons with previous studies.

Panel B shows that the patterns when hourly wages are used as the dependent variable are qualitatively different than those observed when using total earnings or an indicator of positive earnings as the outcome measure. One important difference is that the estimates in Panel B show that the unconditional Black-white wage gap did not substantively change across the three data sets, in contrast to the strongly increasing gap in total earnings gap. Furthermore, the size of the reduction in the Black-white hourly wage gap from controlling for human capital follows an inverted u-shaped pattern: In the NLS-OC the decrease in the wage gap from controlling for human capital is 7.8 log points or 31.2%, in the NLSY-79 the magnitude of this reduction rises to 12.4 log points or 44.4%, but in the NLSY-97 it falls again to 7.9 log points or 37%.[17]

These patterns in hourly wages among working men, while generally different than the patterns for outcomes that also incorporate the extensive margin, are still primarily driven by changes in the returns to education and test scores. Panel B estimates that the association between an additional year of education and hourly wages increased from 4 log points in the NLS-OC to 6.1 log points in the NLSY-79, but then remained stagnant at 6.2 log points in the NLSY-97. Likewise, the association between a standard deviation increase in test scores and hourly wages increased from 3.6 log points in the NLS-OC to 7.6 log points in the NLSY-79, but then decreased to 2.6 log points in the NLSY-97.[18] These patterns highlight that the widely discussed increases in the return to skill in the late 20th Century and early 21st Century apply primarily to the extensive margin.

In addition to hourly wages, relevant intensive earnings margins include the number of weeks worked over the course of the year and the number of hours worked in a typical week, and another potentially informative intensive margin measure is total earnings among men with positive earnings, which simultaneously incorporates hourly wages, hours worked per week, and weeks worked per year. Likewise, there are several distinct extensive margins or reasons for non-work that are of potential intersect, including incarceration, unemployment, and labor force withdrawal. Results for all of these additional outcomes are reported in the online appendix, and similar to the results in Table 6 show that human capital is increasingly important for explaining racial gaps in extensive earnings margins, but that there is no clear positive trend in the share of racial gaps on intensive earnings margins that are explained by human capital.

[17] Neal & Johnson (1996) also use the NLSY-79, and estimate that controlling for AFQT score reduces the Black-white wage gap from 24.4 log points to 7.2 log points, which is a substantially larger reduction than what is shown in Columns 3 and 4 of Table 6. This discrepancy is primarily due to the current paper using a broader sample and controlling for both education and AFQT scores: When I imitate Neal & Johnson (1996) by restricting the sample to men born after 1961 while excluding earlier cohorts and with wages observed in 1990 or 1991, and by controlling only for AFQT rather than AFQT and educational attainment simultaneously, I find that the Black-white wage gap falls from 25.8 log points to 10.9 log points after controlling for test scores, quite similar to Neal & Johnson (1996).

[18] While stagnant or decreasing skill premia from the NLSY-79 ot the NLSY-97 may seem surprising, Castex & Dechter (2014) and Deming (2017) both find that among men the wage returns to test scores fell substantially in the NLSY-97 versus the NLSY-79, and Castex & Dechter (2014) find at most modest increases in the wage returns to education.

3.2 Racial Differences in the Returns to Human Capital

All of the specifications reported above restricted the returns to human capital to be equal across racial groups. Table 7 reports the results of specifications that relax this restriction by regressing total earnings onto educational attainment and test scores separately for the Black and white samples from each survey, and therefore allow the returns to human capital to differ by race.

Table 7 indicates that in all three samples and for both education and test scores, the returns to the skill measures are uniformly larger for Blacks than for whites. For instance in the NLS-OC sample, a one year increase in educational attainment is associated with a 9.1 log point increase in earnings among Blacks and a 7.1 log point increase in earnings among whites, while in the NLSY-79 the estimated returns to a year of education for Blacks and whites are 38.6 log points and 11.9 log points, respectively, and in the NLSY-97 the estimated returns to a year of education for Blacks and whites are respectively 47.1 and 13.3 log points. Similarly, the earnings increase associated with a one standard deviation improvement in test scores is 9.9 log points for Blacks versus 1.7 log points for whites in the NLS-OC, 20.4 log points for Blacks versus 15.2 log points for whites in the NLSY-79, and 80.4 log points for Blacks versus 34 log points for whites in the NLSY-79.

Table 7 also reports P-values from Chi-Square tests of whether the differences in the coefficients for Blacks and whites are statistically significant, and these P-Values indicate that Black-white differences in the returns to human capital often cannot be precisely estimated. In particular in the NLS-OC neither of the differences in coefficients approach being statistically significant at conventional levels, in the NLSY-79 only the difference in the returns to education is statistically significant. In the NLSY-97 both differences are statistically significant. Despite many of the estimated differences being imprecise, the often large magnitudes of the racial differences in the returns to human capital are perhaps surprising, and have a number of potentially important implications.

First, they suggest that higher levels of human capital are strongly rewarded among Black men in post-Civil Rights labor markets, or equivalently that Black men with lower levels of human capital face particularly large earnings penalties.[19] This reduces the likelihood that the lower levels of human capital observed among Black men reflect a conscious choice to not invest in human capital due to anticipated discrimination that would lower the returns to human capital. Rather than reflecting low demand for human capital investment opportunities among African Americans, these differential returns suggest that the supply of available investment opportunities have been inadequate for the Black men in these samples.

The greater returns to human capital among Black men also imply that racial earnings differentials will be most acute at the lower end of the human capital distribution. This is consistent with evidence from Bayer & Charles (2018), who find that much more racial progress has occurred since 1940 among men at the 90th

[19] The returns to human capital are greater among Blacks than whites even when the specifications in Table 7 are estimated with respondents living in the South, where more discrimination might be expected, especially in the NLS-OC sample.

percentile of the earnings distribution than at lower percentiles.[20]

Another question related to the results in Table 7, which is more directly relevant to the current paper's main findings, is how these differential returns affect the conclusion that the contribution of human capital to Black-white gaps in total earnings has grown since the 1960s.

The decompositions estimated above calculated the contribution of human capital to Black-white earnings differences as the product of (1) Black-white gaps in human capital and (2) the association between human capital and earnings. This approach is still applicable in a setting where the association between human capital and earnings varies by race, but the estimates will now depend on which set of coefficients for the human capital characteristics are used. Since this choice is fundamentally arbitrary, I calculate estimates using both the Black and white coefficients, and report the results in the bottom row of Table 7. Specifically, I calculate $(\bar{X_w} - \bar{X_b})'\hat{\beta_b}$ as well as $(\bar{X_w} - \bar{X_b})'\hat{\beta_w}$ in each sample, where $\bar{X_w}$ and $\bar{X_b}$ are the means of the human capital characteristics among whites and Blacks, while $\hat{\beta_w}$ and $\hat{\beta_b}$ are vectors of regression coefficients on the human capital coefficients for whites and Blacks. Note that these quantities are identical to the "explained" portion of a standard Oaxaca-Blinder decomposition, sometimes referred to as "quantities" or "endowment" effects, and which estimate how the gap in a given outcome between two groups would change if those groups had the same average levels of a set of covariates.[21]

The results at the bottom of Table 7 show that when using the coefficients among Blacks, the estimated contribution of human capital to the racial earnings gaps increase from .19 in the NLS-OC, to .53 in the NLSY-79, to 1.19 in the NLSY-97. Conversely, when the white human capital coefficients are used, the estimated contribution of human capital to the racial earnings gap increases from .09 in the NLS-OC, to .26 in the NLSY-79, to .43 in the NLSY-97. The uniformly larger estimates when using the Black coefficients simply reflect the higher returns to human capital among Blacks as shown in the top of Table 7: If the returns to skill are higher, as they are within the Black sample, then equalizing skill levels would be expected to cause a greater reduction in the earnings gap. Most importantly for present purposes, however, is that the estimated contribution of human capital to the Black-white earnings gap grows substantially over time regardless of which skill price estimates are used in the calculation. This occurs because the returns to human capital increase significantly across the three samples for both Blacks and whites, even though the overall returns to human capital are greater for Blacks.

[20] I have also estimated the specification reported in Table 7 using hourly wages as the dependent variable and find higher returns to human capital among Blacks for this outcome as well, although the differences are substantially smaller, suggesting that the differential returns shown in Table 7 are largely due to the extensive margin between work and non-work, rather than higher wages conditional on employment.

[21] Gelbach (2016) discusses the relationship between his decomposition method and the Oaxaca-Blinder decomposition, showing that a simple extension of the Gelbach decomposition nests the traditional Oaxaca-Blinder decomposition.

3.3 Human Capital and Black-White Earnings Disparities Away from the Mean

All of the results presented above focused on comparing the *mean* labor market outcomes of Blacks and whites with varying levels of human capital. A natural extension of this analysis is to estimate the extent to which human capital affects Black-white gaps at other parts of the earnings distribution.

I first note that one of the key findings reported above was that changes in the explanatory power of human capital pertain primarily to the extensive margin between zero and non-zero earnings, and that for a binary outcome of this kind OLS estimates describe the full distribution. The concentration of effects at the extensive margin therefore reduces the practical importance of evaluating changes in human capital's impact at other portions of the earnings distribution.[22] Still, the extent to which human capital's explanatory power differs across the earnings distribution among positive earners, and how any such differences have evolved over time, are of clear interest, and in this section I use the method developed by DiNardo, Fortin & Lemieux (1996) to evaluate effects at other points in the distribution.

The basic approach used by DiNardo, Fortin & Lemieux (1996) is to estimate and apply weights that balance observable characteristics across two groups, and then use the reweighted sample to provide transparent visual characterizations of where in the distribution of an outcome those observable characteristics have the most impact. In the current application the relevant observable characteristics are the human capital measures while the two groups are Black and white men, and kernel density methods can be used to nonparametrically characterize the Black-white earnings gap across full earnings distribution, both with and without applying weights to equalize the human capital levels of Blacks and whites.

For each of the three samples, I first compute the difference in the log annual earnings of the nth percentile white man and the nth percentile Black man, and report the Black-white gap at each percentile. Next, I reweight the white samples in each of the three surveys such that they have similar distributions of education and test scores as the Black samples. Following the procedure suggested by DiNardo, Fortin & Lemieux (1996) I use weights defined by $\frac{P(Black=1|x)P(Black=0)}{P(Black=0|x)P(Black=1)}$, where $P(Black = 1|x)$ denotes the predicted probability that a given respondent is Black given his observed human capital characteristics and is estimated using a probit regression of a Black indicator onto education and test scores, while $P(Black = 1)$ is the unconditional share of each sample that is Black.[23] The Black-white differences in log earnings at each percentile using the reweighted samples are reported alongside the unconditional gaps, and provide a

[22] I also note that median regression has been widely used in the literature to help account for non-participation, since under the reasonable assumption that the potential earnings of non-workers would be below the 50th percentile, quantile regression at the median will accurately estimate the median Black-white earnings gap of the full population. But because the current paper treats zero-earnings as an observed labor market outcome, rather than considering the earnings of non-working men as being unobservable, this type of exercise is not particularly informative in the current context.

[23] $P(Black = 0|x)$ and $P(Black = 0)$ are defined similarly. See DiNardo, Fortin & Lemieux (1996) for detailed derivations of these weights. Note that I reweight the white sample to have human capital characteristics similar to those of the Black sample, and not the reverse. While either approach is valid in principle, reweighting the Black sample to have observables similar to the white sample is likely to encounter a common support problem, since it is more difficult to find Black respondents with human capital levels comparable to those of white respondents. In similar applications Barsky et al. (2002) and Heywood & Parent (2012) also choose to apply weights to the white sample.

counterfactual estimate of what the Black-white earnings gap at each percentile would have been if white workers had human capital characteristics comparable to those of Black workers.

Decomposition results are reported in Figure 3. The solid lines of the figure show kernel-weighted local polynomial regressions of the unadjusted Black-white gaps in each of the three samples, the dashed lines show the counterfactual gaps that occur after applying the reweighting procedure, and the dotted lines show the reduction in the black-white gap at each percentile that occurred after the reweighting.

The unconditional gaps shown with the solid lines in Figure 3 indicate that in all three samples, baseline earnings gaps are largest at the lower percentiles, and then fall as higher portions of the earnings distribution are considered before modestly increasing at the very highest percentiles. This contrasts with some previous evidence that Black-white gaps in hourly wages are often highest at the upper percentiles of the wage distribution, for instance Heywood & Parent (2012), and reflects the use of total earnings in Figure 3 rather than hourly wages.[24]

The difference between the unconditional and counterfactual gaps, shown with dotted lines in Figure 3, indicate that in the NLS-OC the amount of the Black-white earnings gap attributable to human capital is essentially uniform across the distribution, while in the NLSY-79 and NLSY-97 human capital's impacts are largest within approximately the bottom quintile of the earnings distribution, and are then smaller and relatively stable across the reminder of the distribution.

Comparing patterns across the three surveys, Figure 3 shows that the explanatory power of human capital at the lower percentiles increases between the NLS-OC and NLSY-79, but then declines again in the NLSY-97, while above approximately the 20th percentile the explanatory power of human capital is similar across the three data sets. This stands in contrast to the reductions in the mean earnings gap reported in Table 1, which were monotonically and strongly increasing across the three data sets, and more closely resembles the patterns for hourly wages reported in Table 6. The patterns in Figure 3 therefore reinforce the finding that extensive margin effects are the key driver of changes in how human capital affects racial earnings inequality, while also showing that the intensive margin impacts are concentrated among men in approximately the bottom quintile of the earnings distribution.

4 Conclusion

This paper showed that the magnitude of reductions in the Black-white total earnings gap which occur after controlling for human capital have become much larger over time. This increasing explanatory power of human capital was overwhelmingly due to growth in the *returns* to human capital, rather than changes in

[24] Even though zero earners excluded from the samples in Figure 3, the lower percentiles have very low earnings that may reflect modest attachment to the labor force or other extensive margin effects. For example the 5th percentile of the Black earnings distribution is equal to just \$4,988 in the NLS-OC, \$1,977 in the NLSY-79, and \$2,647 in the NSLY-97 (all expressed in 2017 dollars).

racial gaps in formal schooling or test scores, and was strongly concentrated at the extensive margin between work and non-work, rather than intensive margins such as hourly wages.

Many of these findings mirror results reported in previous studies, rather than being strictly novel. For instance Chandra (2001), Ritter & Taylor (2011), and Bayer & Charles (2018) all stress the importance of evaluating racial employment gaps in addition to wages, Card & Lemieux (1994, 1996) and Maloney (1994) emphasize that increasing skill prices will exacerbate Black-white labor market disparities if those disparities substantially reflect productivity differences, and Castex & Dechter (2014) and Altonji, Bharadwaj & Lange (2012) investigate changes in the returns to human capital in the NLSY-79 versus the NLSY-97. While these aspects of my approach and findings are previously established, systematically estimating the contribution of human capital to racial inequality throughout the post-Civil Rights Act period provides several insights that were not readily apparent from the existing literature and which meaningfully revise our previous understanding of how and why racial inequality has evolved over the past 50 years.

First, the aggregate contribution of human capital to differences in the earnings of Black and white men grew steadily over the past 50 years, and has never been larger than it is today. This basic pattern applies to both formal schooling and to standardized test performance. Furthermore, among men engaged with the labor market of the late 1960s and the 1970s, human capital was strongly associated with hourly wages, but was only weakly associated with having non-zero earnings, and as a result the average Black man earned lower wages than his white counterpart but was only moderately less likely to be working. This changed dramatically as the latter half of the 20th century progressed, and in the contemporary US labor market, rather than working in a low-wage job less skilled Black men are now frequently incarcerated, unemployed, or have withdrawn from the labor force during their prime working years.

These qualitative changes mirror key structural developments in the labor market and in US society more generally: Skill-biased technical change and increasing returns to skill, job polarization and declining male labor force participation rates, and the rise of mass incarceration. The current study's findings indicate that these developments dramatically changed the extent and manner in which human capital contributes to racial earnings disparities. Given these changes, policies that reduce Black-white human capital disparities can reasonably be expected to also reduce racial earnings inequality, likely more than ever before. But at the same time, the results strongly suggest that policies which attenuate the relationship between having relatively lower levels of human capital and the probability of non-work, for instance race-neutral criminal justice reforms or sustained tight labor market conditions that facilitate labor force re-entry among less skilled men, may do even more to improve the relative economic well being of African American men.

References

Aguiar, M., Bils, M., Charles, K. K., & Hurst, E. (2017). Leisure luxuries and the labor supply of young men. *NBER Working Paper 23552.*

Altonji, J. G., Bharadwaj, P., & Lange, F. (2012). Changes in the characteristics of American youth: Implications for adult outcomes. *Journal of Labor Economics*, 30(4), 783-828.

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. *Handbook of Labor Economics*, 3, 3143-3259.

Autor, D. H., & Duggan, M. G. (2003). The rise in the disability rolls and the decline in unemployment. *Quarterly Journal of Economics*, 118(1), 157-206.

Autor, D. H., Katz, L. F., & Kearney, M. S. (2008). Trends in US wage inequality: Revising the revisionists. *Review of Economics and Statistics*, 90(2), 300-323.

Bacolod, M., & Hotz, V. J. (2006). Cohort Changes in the Transition from School to Work: Evidence from three NLS surveys. *Economics of Education Review*, 25(4), 351-373.

Bayer, P., & Charles, K. K. (2018). Divergent paths: A new perspective on earnings differences between black and white men since 1940. *Quarterly Journal of Economics,* 133(3), 1459-1501.

Bellemare, M. F., & Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, 82(1), 50-61.

Black, S. E., & Devereux, P. J. (2011). Recent Developments in Intergenerational Mobility. *Handbook of Labor Economics*, 4, 1487-1541.

Burbidge, J. B., Magee, L., & Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401), 123-127.

Card, D., & Lemieux, T. (1996). Wage dispersion, returns to skill, and black-white wage differentials. *Journal of Econometrics*, 74(2), 319-361.

Card, D., & Lemieux, T. (1994). Changing wage structure and black-white wage differentials. *American Economic Review*, 84(2), 29-33.

Carruthers, C. K., & Wanamaker, M. H. (2017). Separate and unequal in the labor market: Human capital and the Jim Crow wage gap. *Journal of Labor Economics*, 35(3), 655-696.

Castex, G., & Kogan Dechter, E. (2014). The changing roles of education and ability in wage determination. *Journal of Labor Economics*, 32(4), 685-710.

Chandra, A. (2000). Labor-market dropouts and the racial wage gap: 1940-1990. *American Economic Review*, 90(2), 333-338.

Chay, K. Y., Guryan, J., & Mazumder, B. (2014). Early life environment and racial inequality in education and earnings in the United States. *NBER Working Paper 20539.*

Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2017). Mobility report cards: The role of colleges in intergenerational mobility. *NBER Working Paper 23618.*

Cunha, F., & Heckman, J. (2016). Decomposing trends in inequality in earnings into forecastable and uncertain components. *Journal of Labor Economics*, 34(S2), S31-S65.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics*, 132(4), 1593-1640.

Derenoncourt, E., & Montialoux, C. (2019). Minimum wages and racial inequality. *Quarterly Journal of Economics*, 136(1), 169-228.

Fryer, F. R. (2011). Racial inequality in the 21st century: The declining significance of discrimination. *Handbook of Labor Economics* (Vol. 4, pp. 855-971)

Gelbach, J. B. (2016). When do covariates matter? And which ones, and how much? *Journal of Labor Economics,* 34(2), 509-543.

Goldin, C. D., & Katz, L. F. (2009). *The race between education and technology*: Harvard University Press.

Goldin, C., & Katz, L. (2020). Extending the race between education and technology. *AEA Papers and Proceedings* (Vol. 110, pp. 347-51).

Heckman, J. J., Lyons, T. M., & Todd, P. E. (2000). Understanding black-white wage differentials, 1960-1990. *American Economic Review,* 90(2), 344-349.

Herriott, R. E., & Kohen, A. I. (1973). On the pooling of mental ability measures from different tests: A pragmatic approach. Unpublished appendix.

Juhn, C. (2003). Labor market dropouts and trends in the wages of black and white men. *ILR Review*, 56(4), 643-662.

Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *Quarterly Journal of Economics,* 107(1), 35-78.

Krueger, A. B. (2017). Where have all the workers gone? An inquiry into the decline of the US labor force participation rate. *Brookings Papers on Economic Activity*, 2017(2), 1.

Lang, K., & Lehmann, J. Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, 50(4), 959-1006.

Lang, K., & Manove, M. (2011). Education and labor market discrimination. *American Economic Review,* 101(4), 1467-1496.

Luo, S. (2020). Understanding the Racial Gaps in the Early Careers of Millennial Men. Working paper.

Maloney, T. N. (1994). Wage compression and wage inequality between black and white males in the United States, 1940-1960. *Journal of Economic History*, 358-381.

Mazumder, B. (2018). Intergenerational Mobility in the US: What We Have Learned from the PSID. Working paper.

Myrdal, G. (1944). *An American dilemma.* In New Tribalisms (pp. 61-72): Springer.

Neal, D., & Rick, A. (2014). The prison boom and the lack of black progress after Smith and Welch (0898-2937). *NBER Working Paper 20283.*

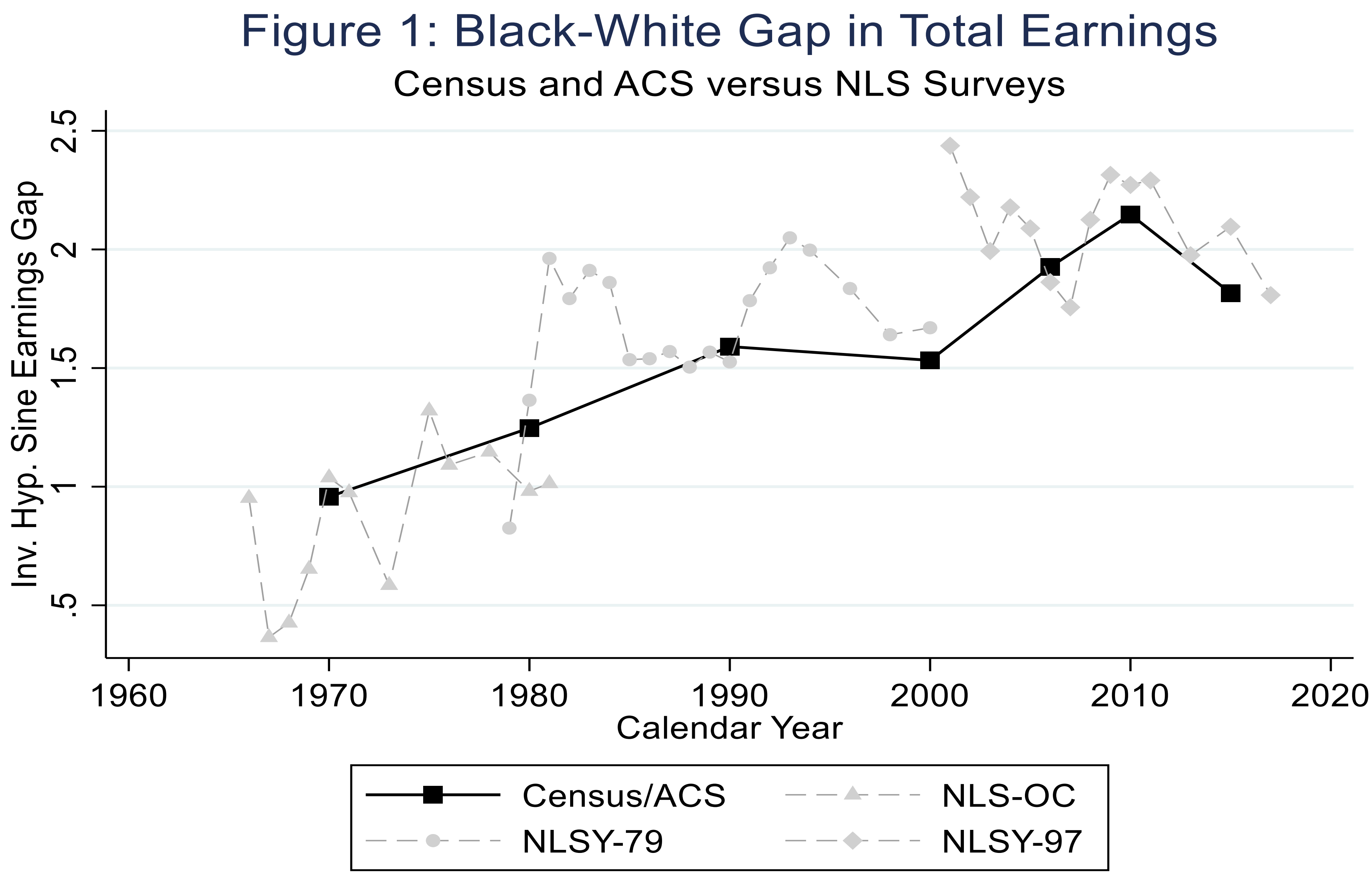
Neal, D. A., & Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. Jo*urnal of Political Economy,* 104(5), 869-895.

Reardon, E. (1997). Demand-side changes and the relative economic progress of black men: 1940-90. *Journal of Human Resources*, 69-97.

Ritter, J. A., & Taylor, L. J. (2011). Racial disparity in unemployment. *Review of Economics and Statistics*, 93(1), 30-42.

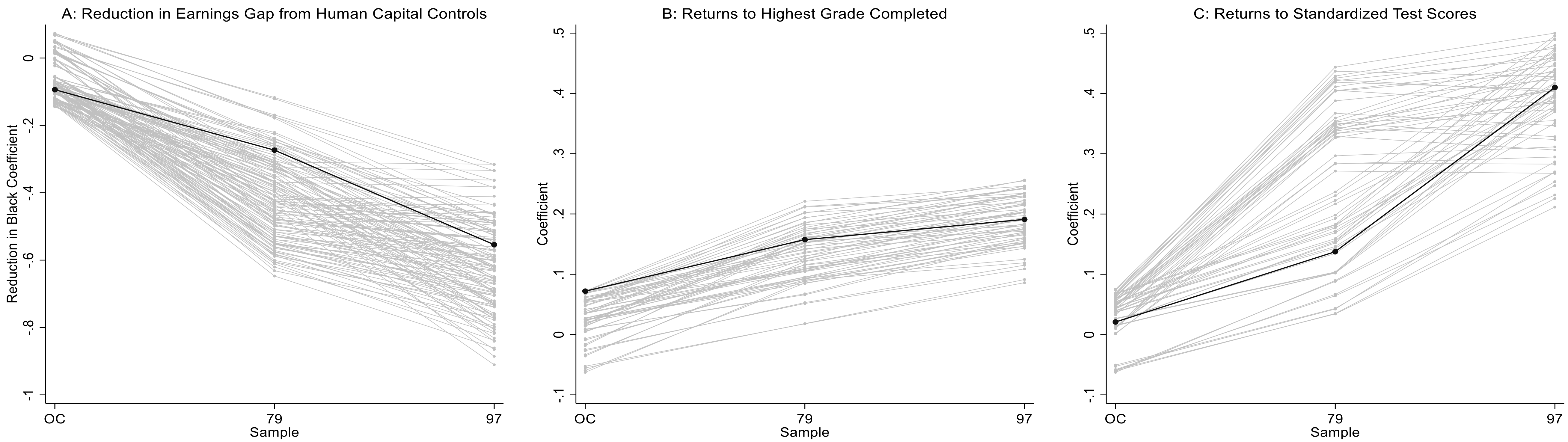
Ruggles, S. et. al (2021). IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021. https://doi.org/10.18128/D01

Western, B. (2002). The impact of incarceration on wage mobility and inequality. *American Sociological Review,* 526-546.



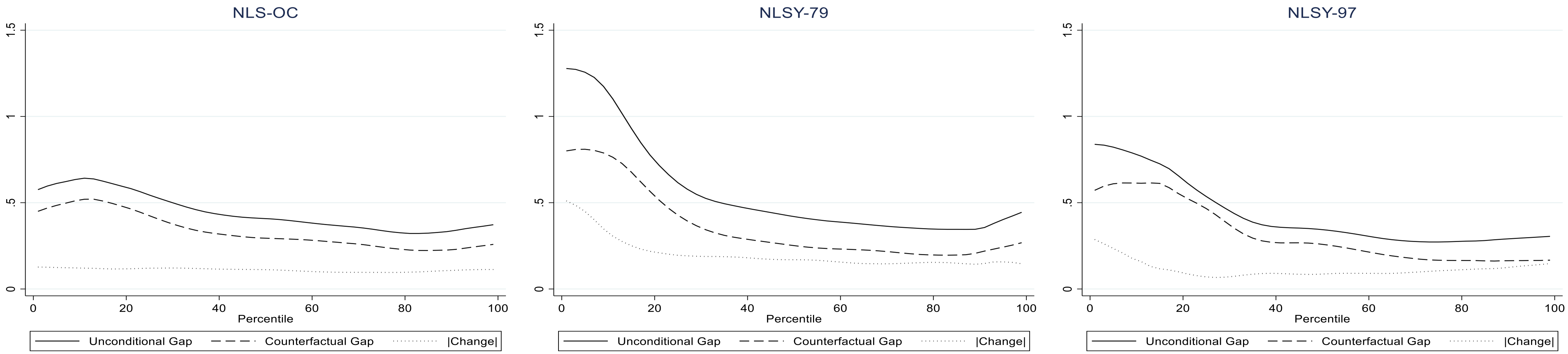
Notes: Figure shows the mean Black-white gap in the inverse hyperbolic sine of total earnings, with zeros included, for men in each National Longitudinal Survey as well as in data from the 1970-2000 Decennial Census samples and the 2007 ACS 3-year sample and the 2012 and 2017 ACS 5-year samples prepared by Ruggles et al (2021). All samples are restricted to non-Hispanic Black and white men aged 21-37 and not currently enrolled in school, and from the 1941-1951 cohorts, the 1957-1964 cohorts, and the 1980-1984 cohorts, which respectively correspond to the cohorts represented in the NLS-OC, NLSY-79 and NLSY-97.

Figure 2: Specification Checks



Notes: The figure reports results of alternative specifications that use each possible permutation of the following modifications to the baseline specification: Using a minimum age of 25 rather than 21, using a maximum age of 32 rather than 37, including current students, measuring human capital with sets of dummies for total years of education completed and standardized test score quartile rather than linear functional forms, including age indicators, a south indicator, and an urban residence indicator as baseline covariates, and applying versus not applying sampling weights. The relevant parameters from the preferred models in Table 1 are shown in bold, while all possible alternative specifications are shown in light gray.

Figure 3: DiNardo, Fortin & Lemieux Decompositions of Black-White Earnings Gap



Notes: The solid lines of the figure show the difference in the inverse hyperbolic sine of annual earnings for the nth percentile white man and the nth percentile Black man, smoothed using kernel-weighted local polynomial regressions. The dashed lines show the counterfactual gaps at each percentile that occur after applying the reweighting procedure described by DiNardo, Fortin & Lemieux (1996). The dotted lines show the difference between the solid and dashed lines, which provides an estimate of the share of the Black-white earnings gap at each percentile attributable to human capital.

Table 1: Unconditional and Conditional Black-White Earnings Differentials

	(1)	(2)	(3)	(4)	(5)	(6)
	NLS-OC		NLSY-79		NLSY-97	
	Baseline	With Controls	Baseline	With Controls	Baseline	With Controls
Black	-0.960***	-0.866***	-1.681***	-1.407***	-1.974***	-1.420***
	(0.144)	(0.151)	(0.100)	(0.111)	(0.135)	(0.136)
Educational Attainment (years)		0.072***		0.157***		0.191***
		(0.018)		(0.021)		(0.021)
Test Score (standard deviations)		0.021		0.137**		0.410***
		(0.045)		(0.059)		(0.063)
Observations	18,138	18,138	42,717	42,717	23,660	23,660
Level Change After Covariates	**-0.09**		**-0.27**		**-0.55**	
Percent Change After Covariates	**-9.8%**		**-16.3%**		**-28.1%**	

Notes: The dependent variable for all models is the inverse hyperbolic sine of total earnings, with zeros included. Observations consist of person-years. All samples are restricted to non-Hispanic Black and white men between the ages of 21 and 37 who are not currently enrolled in school. Sampling weights applied. Standard errors are clustered at the individual level and reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2: Black-White Gaps in Human Capital Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Educational Attainment			Test Scores		
	NLS-OC	NSLY-79	NLSY-97	NLS-OC	NSLY-79	NLSY-97
Black	-1.01***	-0.83***	-1.12***	-1.02***	-1.03***	-0.83***
	(0.13)	(0.08)	(0.11)	(0.07)	(0.03)	(0.04)
Observations	18,138	42,717	23,660	18,138	42,717	23,660

Notes: The dependent variable for Columns 1-3 is educational attainment measured in years, while the dependent variable for Columns 4-6 is standardized test scores measured in standard deviations. Observations consist of person-years. All samples are restricted to non-Hispanic Black and white men between the ages of 21 and 37 who are not currently enrolled in school. Sampling weights applied. Standard errors are clustered at the individual level and reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 3: Decomposition of Black-White Gaps in Total Earnings

	NLS-OC	NLSY-79	NLSY-97
Total Gap Attributable to Human Capital	-0.09**	-0.27***	-0.55***
	(0.04)	(0.05)	(0.06)
Attributable to Education	-0.07***	-0.13***	-0.21***
	(0.02)	(0.02)	(0.03)
Attributable to Test Scores	-0.02	-0.14**	-0.34***
	(0.05)	(0.06)	(0.05)

Notes: Table reports the results of the decomposition procedure described by Gelbach (2016). The first row reports the reduction in the Black coefficient that occurs in each survey after conditioning on educational and test scores, as shown in Table 1. The second and third rows respectively decompose this total reduction into portions attributable to education and test scores. These contributions are calculated as the product of the education (test score) coefficient as reported in Table 1, and the Black-white gap in education (test scores) as reported in Table 2. Observations consist of person-years. All samples are restricted to non-Hispanic Black and white men between the ages of 21 and 37 who are not currently enrolled in school. Sampling weights applied. Standard errors are calculated using the formulas derived in Gelbach (2016) with individual-level clustering and reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 4: Adjusted NLS-OC Results

	(1) Adjusted Coefficients	(2) Adjusted Gaps	(3) Adjusted Contributions	(4) Unadjusted Contributions
Educational Attainment (years)	0.072***	-1.69***	-0.122	-0.073
	(0.018)	(0.139)		
Test Score (standard deviations)	0.029	-1.02***	-0.030	-0.021
	(0.062)	(0.07)		
	Total Contribution of Human Capital		-0.152	-0.094

Notes: This table makes adjustments to account for potential data and measurement issues in the NLS-OC survey. Column 1 reports an adjusted test score coefficient that scales the NLS-OC test score coefficient reported in Table 1 by a first-stage estimate of .731. The educational attainment coefficient in Column 1 reproduces the baseline estimate from Table 1 for reference. Column 2 reports an adjusted Black-white gap in educational attainment that includes NLS-OC respondents without valid test score data, who are less educatead on average. The test score gap reported in Column 2 reproduces the baseline estimate from Table 2 for reference. Column 3 estimates the contribution of education and test scores to the Black-white earnings gap by taking the product of the education/test score coefficients and the Black-white education/test score gaps using the adjusted parameters. Column 4 reproduces the baseline decompositions from Table 3 for reference. Standard errors are clustered at the individual level and reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5: More Educated NSLY-79 and NLSY-97 Samples

	(1)	(2)	(3)	(4)	(5)	(6)
	NLS-OC		NLSY-79		NLSY-97	
	Baseline	With Controls	Baseline	With Controls	Baseline	With Controls
Black	-0.960***	-0.866***	-1.595***	-1.354***	-1.621***	-1.139***
	(0.144)	(0.151)	(0.103)	(0.116)	(0.138)	(0.141)
Educational Attainment (years)		0.072***		0.152***		0.148***
		(0.018)		(0.022)		(0.022)
Test Score (standard deviations)		0.021		0.130**		0.439***
		(0.045)		(0.060)		(0.063)
Observations	18,138		39,135		20,940	
Level Change After Covariates	**-0.09**		**-0.24**		**-0.48**	
Percent Change After Covariates	**-9.8%**		**-15.1%**		**-29.8%**	

Notes: The dependent variable for all models is the inverse hyperbolic sine of total earnings, with zeros included. Observations consist of person-years. All samples are restricted to non-Hispanic Black and white men between the ages of 21 and 37 who are not currently enrolled in school. The NLSY-79 and NLSY-97 samples in this table are further restricted to match the patterns of missing test score data in the NLS-OC. See text for details. Sampling weights applied. Standard errors are clustered at the individual level and reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 6: Black-White Earnings Differentials on the Extensive and Intensive Margin

Panel A: Indicator of Positive Earnings

	NLS-OC		NLSY-79		NLSY-97	
	Baseline	With Controls	Baseline	With Controls	Baseline	With Controls
Black	-0.050***	-0.050***	-0.109***	-0.099***	-0.149***	-0.109***
	(0.012)	(0.013)	(0.008)	(0.009)	(0.011)	(0.012)
Educational Attainment (years)		0.001		0.007***		0.012***
		(0.001)		(0.002)		(0.002)
Test Score (standard deviations)		-0.001		0.004		0.032***
		(0.004)		(0.005)		(0.005)
Observations	18,138		42,717		23,660	
Level Change After Covariates	**0.000**		**-0.010**		**-0.040**	
Percent Change After Covariates	**-0.9%**		**-9.3%**		**-26.5%**	

Panel B: Log Hourly Wages

	NLS-OC		NLSY-79		NLSY-97	
	Baseline	With Controls	Baseline	With Controls	Baseline	With Controls
Black	-0.249***	-0.171***	-0.279***	-0.155***	-0.213***	-0.134***
	(0.024)	(0.023)	(0.014)	(0.016)	(0.021)	(0.022)
Educational Attainment (years)		0.040***		0.061***		0.062***
		(0.004)		(0.004)		(0.004)
Test Score (standard deviations)		0.036***		0.076***		0.026**
		(0.009)		(0.011)		(0.012)
Observations	15,711		38,391		20,540	
Level Change After Covariates	**-0.078**		**-0.124**		**-0.079**	
Percent Change After Covariates	**-31.2%**		**-44.4%**		**-37.0%**	

Notes: The dependent variable is indicated in the subtitle for each panel. Observations consist of person-years. All samples are restricted to non-Hispanic Black and white men between the ages of 21 and 37 who are not currently enrolled in school. Samples in Panel B are further restricted to working men with valid hourly wage data. Sampling weights applied. Standard errors are clustered at the individual level and reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 7: Differences in the Returns to Human Capital by Race

	(1)	(2)	(3)	(4)	(5)	(6)
	NLS-OC		NLSY-79		NLSY-97	
	Black	White	Black	White	Black	White
Educational Attainment (years)	0.091	0.071***	0.386***	0.119***	0.471***	0.133***
	(0.075)	(0.019)	(0.045)	(0.023)	(0.048)	(0.023)
Test Score (standard deviations)	0.099	0.017	0.204	0.152**	0.804***	0.340***
	(0.233)	(0.046)	(0.127)	(0.064)	(0.164)	(0.069)
Observations	2,551	15,587	16,201	26,516	7,877	15,783
P-Value of Test for Difference in Education Coefficients	0.62		0.00		0.00	
P-Value of Test for Difference in Test Score Coefficients	0.48		0.61		0.00	
	$(\bar{x}_w - \bar{x}_b)'\hat{B}_b$	$(\bar{x}_w - \bar{x}_b)'\hat{B}_w$	$(\bar{x}_w - \bar{x}_b)'\hat{B}_b$	$(\bar{x}_w - \bar{x}_b)'\hat{B}_w$	$(\bar{x}_w - \bar{x}_b)'\hat{B}_b$	$(\bar{x}_w - \bar{x}_b)'\hat{B}_w$
Estimated Contribution of Human Capital to Earnings Gap ("Explained Component")	0.19	0.09	0.53	0.26	1.19	0.43

Notes: Column headings indicate the race of the sample used in each model. The dependent variable for all models is the inverse hyperbolic sine of total earnings, with zeros included. Observations consist of person-years. All samples are restricted to men between the ages of 21 and 37 who are not currently enrolled in school. Sampling weights applied. Standard errors are clustered at the individual level and reported in parentheses. P-values for coefficient differences are calculated from Chi-Square tests that cluster at the individual level. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.