Racial Gaps in the Early Careers of Two Cohorts of American Men *

Sai Luo †

Click here for the latest version.

Abstract

This paper studies whether and how early career racial gaps and their underlying forces have changed across two cohorts of young American men (as represented by the NLSY–79 and NLSY–97), with an emphasis on the roles of skills and the school-to-work transition. Tracking Black and white men from early adulthood into their mid-30s and using a semi-parametric decomposition, I estimate the role of different factors contributing to racial labor market gaps within each cohort, and examine how the patterns have changed across cohorts. I establish two main results. First, as racial cognitive skill gaps have narrowed across cohorts, the explanatory power of skills has become smaller in the younger cohort (30%) than in the older cohort (60%). However, unlike the older cohort, the role of the cognitive skills gap in the younger cohort is robust to conditioning on racial differences in family and neighborhood characteristics. In other words, such characteristics are not compensating for skills for younger Black men as well as they did for older Black men. Second, while Blacks in the older cohort narrowed the racial gaps in the school-to-work transition later in their careers, these gaps had a persistent effect (30%) on racial gaps in labor market outcomes for the younger cohort. For the younger cohort, 20% of the persistent impact of racial gaps during the school-to-work transition comes from disadvantages in the location and timing of labor market entry. These findings call for policies that enhance cognitive skills and reducing disparities in access to entry-level jobs.

---

*An earlier draft of this paper was circulated under the title "Understanding the Racial Gaps in the Early Careers of Millennial Men."

†Shanghai University of Finance and Economics (email: luosaiecon@gmail.com). I am grateful for help from Judith Hellerstein, Jessica Goldberg, and Sergio Urzúa. I thank Katharine Abraham, Melissa Kearney, Peter Bergman, Richard Murphy, Eric Humphries, Ben Zou, Seth Murray, seminar participants at the BLS, U Maryland, SJTU-Antai, PAA, ASSA, AEFP, MEA/SOLE, ES Asian meeting, and ES China meeting for helpful discussions and comments. Jennifer Cassidy-Gilbert from the Bureau of Labor Statistics provided excellent assistance on the geocode data. The NLSY geocode data can be obtained by filing an application directly with the Bureau of Labor Statistics (https://www.bls.gov/nls/geocodeapp.htm). This project is supported by funding from the Economic Club of Washington, D.C. and the Maryland Population Research Center.

1 Introduction

Since the Civil Rights Act of 1964, substantial progress has been made in closing racial gaps in various aspects of U.S. society (Smith and Welch, 1989). However, recent evidence casts doubt on whether this progress is continuing (Wilson and Rodgers, 2016; Daly, Hobijn, and Pedtke, 2017; Bayer and Charles, 2018; Chetty et al., 2020; Thompson, 2021). Despite the growing attention and interest from both the public and academia, there is relatively limited evidence on what underlies the observed racial gaps for young men in today's U.S. labor market and whether and how the underlying forces have evolved across cohorts. In fact, the current narrative of the causes of racial labor market gaps is largely shaped by data over two decades old.

Relying on data from previous cohorts, education and skills have been shown to be crucial in understanding racial gaps in labor market outcomes (e.g., Neal and Johnson, 1996; Heckman, Stixrud, and Urzúa, 2006; Urzúa, 2008; Lang and Manove, 2011). Yet, recent evidence suggests that the importance of cognitive skills in the labor market has declined over time (Castex and Dechter, 2014; Deming, 2017) while other factors, such as neighborhood-level influences, seem to have a persistent impact on shaping racial disparities in labor market outcomes (Chetty et al., 2020). Do Black-white differences in education and skills still play an important role in explaining racial gaps for young men in more recent cohorts, as has been found in previous cohorts?

Young men in today's U.S. labor market experienced an important part of their early careers under the Great Recession, and anecdotal evidence shows that they have struggled to get a foothold in the labor market (The Atlantic, 2015; Forbes, 2016). Did Black men in the younger cohort have a worse school-to-work transition than their white counterparts? If so, does this Black disadvantage have a persistent impact on racial labor market gaps observed in later career years, as suggested by past studies (e.g., Kahn, 2010; Schwandt and Wachter, 2019)? Different answers to these questions suggest different potential policy pathways to reduce racial gaps in the labor market.

In this paper, I provide some of the first evidence on whether and how the Black-white labor market gaps and the underlying forces have changed across two cohorts of young American men. In particular, I focus on the dual roles of pre-market skills (including education) and the school-to-work transition. This analysis is facilitated by using two similarly constructed samples of the National Longitudinal Survey of Youth, the NLSY–79 and the NLSY–97, which are nationally representative samples of young Americans born in 1957–1964 and 1980–1984, respectively.[1] A large body of research has used the NLSY–79 data to study racial gaps in the older cohort (e.g., Neal and Johnson, 1996; Altonji and Blank, 1999; Heckman, Stixrud, and Urzúa, 2006), and my paper adds to this literature by focusing on what has and has not changed in the younger cohort of the NLSY–97.

Tracking the two cohorts of Black and white non-Hispanic men from early adulthood to their mid-30s, I first document the Black-white differences in employment and earnings trajectories and how the pattern has changed across cohorts. I find that, over the first eight years after the completion of schooling, the racial gap in employment and earnings narrows significantly in the older cohort (NLSY–79) but stays broadly steady in the younger cohort (NLSY–97). In other words, the initial racial labor market gap is much more persistent through the early career years in the younger cohort.

In my main analysis, I use a semi-parametric decomposition method (DiNardo, Fortin, and Lemieux, 1996) to examine the explanatory power of pre-market skills and the school-to-work transition with respect to the observed racial gaps in employment and earnings six to eight years after schooling completion, when the labor market outcomes reached a relatively stable stage. My measures of skills include highest grade completed, cognitive skills, non-cognitive skills, and social skills. My measures of the school-to-work transition include employment outcomes in the first year post-schooling and local labor market conditions (unemployment rates) upon schooling completion.[2] I also incorporate measures of family background and childhood neighborhood characteristics to control for racial differences at family and neighborhood levels.[3]

I establish two main empirical results regarding the roles of skills and the school-to-work transition. First, education and skills play a key role in explaining racial labor market gaps in the younger cohort, a result that has been documented in the older cohort (NLSY–79) as well. In the NLSY–97, racial differences in measured education and skills explain about 30% of

[1] Throughout this paper, I refer to the NLSY–79 as the "older" cohort and the NLSY–97 as the "younger" cohort. Altonji, Bharadwaj, and Lange (2012) is one of the first papers that compares the NLSY–79 and the NLSY–97 cohorts, and their focus is on how the characteristics of young Americans have changed across cohorts. As a recent example, Thompson (2021) combines the two NLSY cohorts with an older cohort (NLS-Young Men), and examines how the racial gaps have evolved from a longer time perspective.

[2] Unemployment rates data are taken from the Local Area Unemployment Statistics (LAUS) program.

[3] To keep the decomposition results comparable between the two cohorts, I mainly follow Altonji, Bharadwaj, and Lange (2012) and include the individual, family, and neighborhood variables that are similarly constructed or can be appropriately conformed. Childhood neighborhoods are identified using the NLSY restricted geocode files. I discuss variable definitions in detail in the next section.

the racial gaps in employment and earnings. The explanatory power of education and skills is attributable primarily to racial differences in cognitive skills, as measured by the Armed Forces Qualification Test (AFQT) score, rather than to differences in formal schooling or social and non-cognitive skills.

In the older cohort, racial differences in education and skills explain about 60% of the racial labor market gaps, which is consistent with past studies using the NLSY–79 data (Neal and Johnson, 1996; Urzúa, 2008).[4] From the NLSY–79 to the NLSY–97 cohort, the Black-white gap in measured cognitive skills (AFQT score) has narrowed by about 30%. The cross-cohort change in the explanatory power of education and skills can be largely attributed to this falling racial gap in cognitive skills.

What underlies the racial skill gaps? The AFQT score, for example, is observed at ages 12–18 and can be a function of family investments and neighborhood influences in early childhood.[5] To shed light on this question, I estimate the contribution of education and skills after conditioning on measured racial differences in family background and childhood neighborhood characteristics. In the DiNardo-Fortin-Lemieux decomposition, this is intuitively equivalent to first residualizing the racial skill gaps by the observed racial differences in family and neighborhood characteristics, and then evaluating the contribution of this residualized racial skill gaps to the observed racial labor market gaps.[6]

In the younger cohort, the explanatory power of education and skills remains quantitatively robust (about 30%) even after conditioning on family and neighborhood characteristics. Conversely, in the older cohort, the explanatory power of education and skills disappears almost entirely after accounting for racial differences in family and neighborhood characteristics. This suggests that, although the racial skill gaps have narrowed in the younger cohort, especially with respect to cognitive skills, the remaining racial skills gaps seem to have originated from sources not easily measured by the family and neighborhood variables in the NLSY.

Persistent racial discrimination is one potential explanation of persistent racial skills gaps (Donohue and Heckman, 1991; Pager, 2003; Bertrand and Mullainathan, 2004; Charles and

[4] In Neal and Johnson's (1996) influential work, cognitive skills (AFQT score) account for about 60% of the racial wage gap between Black and white men. Urzúa (2008) makes a distinction between measured cognitive skills (AFQT score) and underlying cognitive ability and shows that cognitive ability explains about 40% of the racial gaps in wages and earnings.

[5] The impact of family and neighborhood (including schools) in the skill accumulation process is extensively discussed in the literature (e.g., Cunha et al., 2006).

[6] See Section 4 and Appendix B for methodological details.

Guryan, 2008). Black men and their parents who anticipate there will be labor market discrimination may underinvest in skill accumulation.[7] While I do not investigate in detail the origins of the racial skill gaps in the younger cohort, my findings reinforce the literature that emphasizes the critical role of skill development (Heckman, Stixrud, and Urzúa, 2006; García et al., 2020).

The second empirical result concerns the school-to-work transition. Although Black men in both cohorts started their careers with significant disadvantages compared to their white counterparts, the school-to-work transition has a particularly important role in explaining racial gaps in later labor market outcomes for the younger cohort. To isolate the impact of the racial gap in the school-to-work transition, I estimate the contributions of the transition after conditioning on racial differences in skills, family, and neighborhood characteristics.[8]

In the NLSY–97 cohort, the racial gap in the school-to-work transition explains up to 30% of the racial labor market gaps observed over the sixth to eighth years post-schooling. Of the two transition measures, the racial gap in the local unemployment rates upon schooling completion explains about 20% of the racial labor market gaps in the NLSY–97. This finding is intuitively consistent with the series of studies showing that poor outcomes early in the school-to-work transition have a long-lasting impact on future labor market outcomes (Kahn, 2010; Schwandt and Wachter, 2019).

In contrast, for the NLSY–79, the school-to-work transition (conditioned as above) does not explain the racial gaps in later outcomes. What makes the school-to-work transition in the younger cohort different? One possible explanation is the Great Recession, which covered a large part of the early career years for the NLSY–97 cohort. The inability to make a successful school-to-work transition during the Great Recession could have had a greater effect on future labor market prospects, compared to smaller recessions experienced by young men from the NLSY–79 cohort.[9]

The remainder of this paper proceeds as follows. Section 2 describes the NLSY datasets and how I create the concordance of variables between the two cohorts. Section 3 presents the early

[7] The impact of discrimination can go beyond the labor market. For example, racial discrimination in the criminal justice system reduces the labor market prospects of Black men, which could further discourage Black children and families from investing in education and skills. Racial discrimination in the housing market and in the education system could limit the opportunities for Black children to live and learn in promising environments and therefore restrict the possibility of narrowing the racial skill gap.

[8] After conditioning on racial differences in the pre-market factors, this decomposition uses the residualized (not raw) racial gap in transition. See Section 4 and Appendix B for methodological details.

[9] Another potential explanation is discrimination. The racial employment gap in the first year post-schooling could be picking up discrimination faced by Black men in their initial labor market experiences.

career labor market trajectories for Black and white men in the two NLSY cohorts. It also shows how the racial differences observed in the pre-market characteristics have changed across the two cohorts. Section 4 introduces the semi-parametric decomposition method. Section 5 discusses the decomposition results, with an emphasis on the dual roles of skills and the school-to-work transition. Section 6 concludes.

2 Comparing Two Cohorts of Young American Men

This paper examines whether and how racial labor market gaps and their underlying explanatory factors have evolved across cohorts. I mainly rely on the 1979 and 1997 cohorts of the National Longitudinal Survey of Youths (NLSY–79 and NLSY–97). With proper sample weights, the NLSY–79 and the NLSY–97 are nationally representative of young Americans born between 1957–1964 and 1980–1984, respectively. In this section, I discuss the advantages of the NLSY datasets and how I construct the samples and variables.

2.1 Data: NLSY–79 and NLSY–97

My analysis uses Black and white non-Hispanic men from both the main sample and the minority subsample of the NLSY–79 and the NLSY–97.[10] The NLSY dataset suits my analysis in four ways. First, it includes a retrospective monthly record of school enrollment and a retrospective weekly record of work status, which I use to define the exact time at which a young man completes schooling and to track his employment and earnings outcomes year by year.[11] My main analysis uses a balanced panel of young men who have not been enrolled in school for at least eight years (or 96 months) and tracks their labor market outcomes through the first eight years post-schooling. As I show below, the work trajectories of both Black and white men in both cohorts reach a relatively steady stage about six to eight years beyond school completion.

[10] I do not include the economically disadvantaged white subsample or the military subsample of the NLSY–79. I exclude men of other races (Asian, mixed, etc.). For brevity, I refer to white non-Hispanic men as white.

[11] I define schooling completion following the literature (Light and McGarry, 1998; Neumark, 2002). Specifically, I identify the first month when a young man was no longer enrolled in school and define the next 12 months as the first year post-schooling. My findings are robust if I define the first post-schooling year as the first calendar year that a young man is completely out of school. If a young man graduated from high school, worked for a few years, went back to college, and rejoined the workforce later, his post-schooling experiences are defined to include only the post-college years. This definition therefore excludes two kinds of work experience: 1) part-time jobs while enrolled in school and 2) relatively temporary work spells that are followed by returning to school (as in the previous example).

Second, the NLSY records rich information on individual, family, and neighborhood characteristics, all of which are critically important for my decomposition analysis. In particular, it includes a measure of cognitive skills (AFQT score) that has been shown by past studies to be a key determinant in understanding racial gaps in the U.S. labor market (Neal and Johnson, 1996).[12] Third, the NLSY follows incarcerated respondents. This is extremely important in the context of understanding racial gaps in labor market outcomes, because Black men are overrepresented in the incarcerated population and much of the NLSY–97 cohort has grown up under a historically high incarceration rate.[13] Bayer and Charles (2018) show that ignoring the ever-increasing prison population leads to an understatement of the racial earnings gap, especially since the late 1970s.

Last, and most importantly, the NLSY–79 and the NLSY–97 surveys are designed and administered in a similar way, so that many of the key variables from the two cohorts are comparable either directly or after some concordance, facilitating a valid comparison between the two cohorts.[14] In my main analysis, I use the individual and family variables constructed by Altonji, Bharadwaj, and Lange (2012) and Deming (2017) and create measures of neighborhood characteristics using restricted-use geocode files.

2.2 Sample Decisions and Variable Definitions

To make sure that the early career trajectories are comparable between the two NLSY cohorts, I construct the samples following two principles. First, the school enrollment record starts in 1980 for the NLSY–79 cohort and in 1997 for the NLSY–97 cohort. For young men who completed schooling before the enrollment record started, the exact school exit time cannot be identified. To minimize this issue without losing too much sample size, I exclude NLSY–79 respondents who were older than 18 as of 1980.[15] For both cohorts, I also exclude young men who were already out of school when the enrollment record started or were still enrolled in school as of

[12] The AFQT score is constructed based on multiple sections of the Armed Services Vocational Aptitude Battery (ASVAB), a set of tests developed by the Department of Defense (DOD) for screening military enlistees and assigning them to military occupations. Economists have long used the AFQT score, as well as other tests in the ASVAB, to measure skills and abilities (e.g., Neal and Johnson, 1996; Heckman, Stixrud, and Urzúa, 2006; Prada and Urzúa, 2017). However, it is important to note that only about 80% of the NLSY–97 respondents took the ASVAB. See Appendix Table A.1 for how sample size changes with the inclusion of different variables.

[13] The incarceration rate has more than tripled since 1980, and criminal justice policies have been shifting toward more punitive treatment, the burden of which falls disproportionately on Black men (Neal and Rick, 2014; Council of Economic Advisors, 2016).

[14] The schooling and work history records are also similarly constructed for the two NLSY cohorts.

[15] All NLSY–97 respondents were younger than 18 as of 1997.

the latest survey wave that I use for that cohort.[16]

Second, as of the most recent survey wave for the NLSY–97 cohort (in 2015), the respondents were around ages 30–34. I focus my analysis of the NLSY–79 cohort on survey years 1979–1996, so that, in the last survey wave that I use for the older cohort (1996), those respondents were in an age range (31–34) close to the NLSY–97 cohort. To keep the restricted samples nationally representative, I apply the custom sample weights created by the Bureau of Labor Statistics.

I construct measures of education and skills and family background following the literature. Specifically, my measures of education and skills include four variables: highest grade completed, AFQT score (as a measure of cognitive skills), non-cognitive test score, and social test score. The AFQT score is measured at different ages for the two NLSY cohorts and for people in the same cohort (ages 15–23 in the NLSY–79 cohort and ages 12–18 in the NLSY–97 cohort). The test format also changed from a paper-based test in the NLSY–79 to a computer-based adaptive test in the NLSY–97. Altonji, Bharadwaj, and Lange (2012) carefully adjust for different test-taking ages and test format changes between the two cohorts, and I use their adjusted AFQT score.[17]

Unlike the AFQT score for cognitive skills, there is no consistent measure of non-cognitive or social skills for the NLSY–79 and NLSY–97 cohorts. Deming (2017) selects survey questions and/or tests from the two cohorts that seem to measure similar skills and creates standardized non-cognitive and social test scores. Without a better way to handle this issue, I use the test scores from Deming (2017). It is important to note that my decomposition results are quantitatively robust with or without including the non-cognitive and social test scores.

My family background variables include three variables constructed by Altonji, Bharadwaj, and Lange (2012): parental income measured at the first wave of each cohort, mother's highest grade completed, and family structure (whether the respondent lives with both parents) during

[16] In other words, my sample includes young men who completed schooling after the enrollment record started and before the latest survey wave that I use for that cohort.

[17] An important question related to the AFQT score, like almost all other psychometric test scores for skills and abilities, is whether the test is biased in favor of one group over another. Since its first introduction by the Department of Defense for screening enlistees and assigning them to different occupations, a key question especially relevant for this paper is whether the AFQT score is racially biased. In 1991, the National Academy of Sciences (NAS) led a study in the military focusing on the test's racial fairness and concluded that the AFQT score does not systematically underpredict the job performance of Blacks relative to whites (Wigdor and Green Jr., 1991). The NAS study provides the best evidence to date regarding the test's fairness, as it directly observes and measures military job performance and links it to the AFQT score, which is not generally available in civilian datasets. Whether the findings of the NAS study can be applied to the civilian population is an open question; some studies cast doubt on the score's racial fairness (Rodgers and Spriggs, 1996), while others conclude otherwise (Heckman, 1998).

childhood.[18]

I construct measures of childhood neighborhood characteristics using the restricted-use geocode files for the NLSY. For the NLSY–79 cohort, I link county of residence at age 14 with county socioeconomic conditions created from the 1980 Census; for the NLSY–97 cohort, I link county of residence at age 12 with the 2000 Census. The socioeconomic variables include county population, median household income, poverty rate, and the share of men with a college education.[19] I also include the same variables but at the state level. To capture some of the within-county variations in neighborhood quality, I further account for whether childhood residence is in a central city, whether it is in a metropolitan statistical area (MSA), and whether it is in an urban or rural area.[20]

I construct two different measures of the school-to-work transition. The first measure is a flexible function of weeks worked in the first year post-schooling. Specifically, I include a series of indicator variables for the number of weeks worked in the first year post-schooling: 1–9 weeks, 10–19 weeks, ..., 40–49 weeks, and 50 weeks or more. The second measure uses the geocode files to link a young man's state of residence in the year that he completed schooling (and entered the labor market) with state average unemployment rates from the Local Area Unemployment Statistics (LAUS) program. This provides a different and presumably more exogenous measure of the transition. For the younger cohort, for which LAUS has more detailed data at the *county* level, I also construct unemployment rates at entry county-year as another measure of transition.[21]

The labor market outcomes of interest are employment and earnings. My employment measures are constructed from the retrospective weekly work records discussed above. Specifically, I summarize information from the weekly level to the yearly level for every year post-schooling, and then compute a set of employment measures, including the number of weeks worked per

[18] Family structure is measured at age 14 for the NLSY–79 cohort and at the first wave (ages 12–16) for the NLSY–97 cohort.

[19] Residence at age 14 is reported for the NLSY–79, and residence at age 12 is reported for the NLSY–97. This difference will not invalidate comparisons between the two cohorts (for the purpose of constructing neighborhood measures) if residence does not change much from age 12 to 14. In the geocode files of the NLSY–97, residence at the time of the first survey (in 1997, when the respondents were 12–16) is also reported. As suggestive evidence, I find that the state of residence does not change for 96% of the NLSY–97 respondents from age 12 to the time of the first survey, and county of residence does not change for 93% of the NLSY–97 respondents over the same time period.

[20] This information is measured at the first wave (ages 14–17) for the NLSY–79 cohort and at age 12 for the NLSY–97 cohort.

[21] County-level unemployment rate data in LAUS only go back to 1990 and therefore are unavailable for most of the NLSY–79 sample years.

year, any employment, half-year employment (worked for at least 26 weeks), and full-year employment (worked for at least 50 weeks). In the decomposition, I focus on the number of weeks worked per year as the employment measure. My earnings measure is annual earnings reported in each survey wave. I adjust annual earnings in both cohorts to 2013 dollars, and use inverse hyperbolic sine to allow for zero earnings. For simplicity, I use the word "log" instead of inverse hyperbolic sine throughout the paper.

The final sample includes 444 white men and 271 Black men from the NLSY–79 cohort and 825 white men and 396 Black men from the NLSY–97 cohort. These young men have completed schooling for at least eight full years and have a complete list of the aforementioned variables of education and skills, family background, and childhood neighborhood characteristics. In Appendix Table A.1, I present how the sample size changes with each of my sample and variable choices. In Appendix Table A.2, I present summary statistics for key individual and family variables in both the original NLSY sample and my final sample. Although the original NLSY sample is much larger (3,874 and 3,455 Black and white non-Hispanic men in the NLSY–79 and NLSY–97 respectively), there are only modest differences from my final sample along the measured individual and family characteristics.[22]

3 Documenting Cross-Cohort Changes in Racial Gaps

In this section, I first show how the racial gaps in early career trajectories have evolved across the two NLSY cohorts, where I trace employment and earnings outcomes from the first year to the eighth year after a young man completed schooling. I then provide descriptive evidence on how the racial gaps in skills, family, and neighborhood characteristics have changed across cohorts.

[22]For the NLSY–79 cohort, both Black and white men in my final sample have slightly higher AFQT score and parental income than the original sample. Individual education, mother's education, and family structure are almost the same between my final sample and the original sample for the NLSY–79. For the NLSY–97 cohort, both Black and white men in my final sample have lower educational attainment and AFQT scores than the original sample. This is partly due to my sample construction process, which is more likely to exclude Black and white men who go to college and/or graduate school (and therefore do not have eight consecutive years of out-of-school experiences by 2015). Parental income, mother's education, and family structure are almost the same between my sample and the original sample for the NLSY–97. See Table A.2 for details.

3.1 Racial Gaps in Early Career Trajectories

Table 1 summarizes the early career outcomes in two periods: the school-to-work transition stage, defined as the first year post-schooling, and the later stage, defined as the sixth to eighth years. I specifically look at the sixth to eighth years because this is when employment and earnings outcomes of young men reached a relatively stable stage. I also summarize the growth in the outcomes from the first year to the sixth to eighth years.

In both cohorts, Black men fell substantially behind their white counterparts during the transition stage, along multiple margins of employment and earnings. In the NLSY–79 cohort, it took Black men 30 more weeks to get their first job; in the NLSY–97 cohort, it took them 22 more weeks. The decrease across cohorts in the racial gap is quantitatively large although statistically insignificant, partly due to the sample size. In the first year post-schooling of both cohorts, Black men were less likely than white men to have any job, to work for half a year ($\geq$ 26 weeks), or to work for a full year ($\geq$ 50 weeks). In the NLSY–79 cohort, Black men worked 13 fewer weeks in the first year, and they worked 9 fewer weeks in the NLSY–97 cohort. The racial gaps in all of these employment outcomes have fallen across cohorts, and some of the declines (any employment, full-year employment) can be statistically distinguished from zero. The racial gap in annual earnings is large and significant in both cohorts, but the change from the NLSY–79 cohort to the NLSY–97 cohort is minimal.

If one only looks at the racial gaps at the transition stage, a natural conclusion is that the relative position of Black men has actually improved across cohorts, especially in employment outcomes.[23] However, the longitudinal structure of the NLSY data allows me to track Black and white men through their early career years and ask: Have these racial gaps converged or persisted through the sixth to eighth years, and how have the trends changed across cohorts? For Black and white men in both cohorts, weeks worked per year and annual earnings increased from the transition stage to the sixth to eighth years, with the increase greater for Black men (i.e., they caught up to white men). As a result, the racial gaps in weeks worked and earnings diminished over the first eight years post-schooling. In the NLSY–79 cohort, the racial gap in weeks worked per year fell from 13 weeks in the first year to 6 weeks, on average, in the sixth

[23] Black men in my sample are slightly younger than their white counterparts upon school completion, which is a mechanical result of the Black-white education gap (as shown in Table 2). The age gap at school completion is small in magnitude for both cohorts (about half a year for the NLSY–79 and about three months for the NLSY–97) and the cross-cohort change is small and statistically insignificant.

to eighth years. In the NLSY–97 cohort, the racial gap in weeks worked per year fell from 9 weeks to 7 weeks. The convergence in weeks worked between Black and white men is smaller in the younger cohort than in the older cohort (by about 5 weeks per year), and the cross-cohort difference is statistically significant. The convergence in earnings is also smaller in the younger cohort but is insignificant.

For a large share of young men in the NLSY–97 sample, their early career years overlapped with the Great Recession, as shown in Appendix Figure A.1. Because less-advantaged groups have been shown to experience more economic losses during economic downturns (e.g., Schwandt and Wachter, 2019), it is plausible that the Great Recession suppressed the potential for this younger cohort of Black men to catch up in labor market outcomes in their first few years out of school.

In addition to the summary statistics, Figures 1–2 directly plot the trajectories of various employment and earnings outcomes through the first eight complete years post-schooling, allowing a visual examination of how the *shapes* of employment and earnings trajectories have changed between the two cohorts.[24] In the NLSY–79 cohort, young men of both races experienced clear upward-sloping career trajectories, as their employment and earnings outcomes gradually improved, especially in the first four to five years after completing schooling. Importantly, the upward-sloping trend is stronger for Black men, generating the catching-up shown in Table 1. In the NLSY–97 cohort, however, the employment and earnings outcomes of both races either stayed largely stable during the first eight years post-schooling or experienced flatter upward-sloping trajectories than young men in the NLSY–79 cohort. This latter evidence is consistent with the anecdotal observation that younger Americans have struggled to gain a foothold in the labor market and to climb up the career ladder (The Atlantic, 2015; Forbes, 2016).

Another important pattern from Figures 1–2 is that the employment and earnings trajectories had more fluctuations and steeper growth in the first few years and started to enter a relatively stable stage around the fourth and fifth years. This is consistent with historical evidence that most job mobility and wage growth happens in the first few years of one's career (e.g., Topel and Ward, 1992). Therefore, in the decomposition analysis in the next section, I focus on racial gaps in employment and earnings outcomes measured over the sixth to eighth years post-schooling.

[24]More outcomes are presented in Appendix Figure A.2.

3.2 Racial Gaps in Pre-Market Factors: Education and Skills, Family Background, and Childhood Neighborhood

To examine how much of the observed racial gaps in labor market outcomes can be explained by underlying characteristics, it is important to first understand whether there are racial gaps in these characteristics and whether and how such gaps have changed across cohorts. Table 2 compares Black and white men in each cohort along the series of pre-market characteristics (education and skills, family background, and childhood neighborhood) discussed in the previous section and tests whether the racial gap in each factor has changed significantly across cohorts.[25]

Among the four variables measuring education and skills, the racial gaps in highest grade completed and AFQT score percentile are statistically significant in both cohorts, and the racial gap in social test scores is statistically significant only in the NLSY–97 cohort. From the NLSY–79 cohort to the NLSY–97 cohort, the starkest change is a decrease in the racial gap in AFQT score. On average, the racial AFQT score gap has fallen significantly, by more than 10 percentiles (which is about 30%). This observation is consistent with Altonji, Bharadwaj, and Lange (2012), who find that the skill-related characteristics of young Americans have changed dramatically between the two NLSY cohorts and that the cross-cohort change is largely driven by AFQT score. The racial gap in the social test score has increased significantly, but the magnitude of change is arguably modest (about 0.35 standard deviations). The racial gap increases slightly in the highest grade completed and decreases slightly in non-cognitive test scores. Both changes are indistinguishable from zero.

For family background characteristics, the racial gaps in all three variables are statistically significant within each NLSY cohort. Comparing across cohorts, the racial gap in parental income has increased significantly, while the racial gap in mother's education has fallen (but the latter change is not statistically significant). Young men of both races are less likely to grow up in a two-parent family in the NLSY–97 cohort than in the NLSY–79 cohort, but the racial gap in childhood family structure stays almost unchanged between the two cohorts.

In both cohorts, Black men tend to grow up in counties and states with a larger population,

[25] Altonji, Bharadwaj, and Lange (2012) create an index of skills for young Americans of the two NLSY cohorts and show that the racial skill gap has narrowed, on average, between Black and white men from the NLSY–79 cohort to the NLSY–97 cohort. The authors construct the skill index based on a set of skill measures (including schooling, AFQT score, parental education, family structure, and school-to-work transition measures) and its relationship with wages in the NLSY–79. The authors also show how the skill distribution has changed across cohorts.

lower median household income, higher poverty rate, and lower share of men with a college education. Some of these racial gaps in neighborhood socioeconomic conditions (such as county median household income and poverty rate) appear to have narrowed between the NLSY–79 cohort and the NLSY–97 cohort. During their childhood, Black men are more likely to live in central cities and white men are more likely to live in suburban areas (MSA, non-central city, urban areas). This racial difference seems to have also decreased across cohorts.

4 Decomposition Method

To assess how the different underlying factors contribute to the documented racial employment and earnings gaps, I rely on the semi-parametric decomposition method introduced by DiNardo, Fortin, and Lemieux (1996, hereinafter DFL).[26] This method relaxes the parametric functional forms imposed by the classical Oaxaca-Blinder decomposition (or other decomposition methods based on linear regressions, such as Gelbach (2016)) on the relationship between labor market outcomes (such as employment and earnings) and underlying characteristics.[27]

In this section, I briefly describe the DFL method in the context of understanding racial labor market gaps.[28] In a nutshell, the DFL decomposition constructs the counterfactual distribution of labor market outcomes that can be used to answer questions such as "What earnings would white men have had if they had the same underlying characteristics (such as education and skills, family background, childhood neighborhood, or the school-to-work transition) as Black men in the same cohort?" The difference between the actual and counterfactual earnings for white men can then be seen as the contribution of Black-white differences in the underlying characteristics to the Black-white earnings gap.[29]

[26]Altonji, Bharadwaj, and Lange (2012) apply the DFL method to study how the characteristics of young Americans have changed from the NLSY–79 cohort to the NLSY–97 cohort and what it means for the labor market prospects of the NLSY–97 cohort.

[27]Some influential studies that focus on understanding labor market racial gaps have relied on regression-based estimates, which impose strong assumptions on parametric (mostly linear) functional forms (Neal and Johnson, 1996; Chetty et al., 2020). However, there is evidence that some of the parametric assumptions widely imposed in classical regression specifications are not supported by the data (Heckman, Stixrud, and Urzúa, 2006). In the context of racial wealth gaps, Barsky et al. (2002) show that the classical Oaxaca-Blinder decomposition (Blinder, 1973; Oaxaca, 1973), which imposes strong functional form assumptions, results in misleading conclusions regarding the explanatory power of racial gaps in earnings on racial wealth gaps.

[28]I leave the methodological details and a discussion of how DFL estimates relate to other estimates in the literature to Appendix B.

[29]The DFL method constructs the counterfactual outcomes for white men by reweighting the white men sample to match the Black men sample in one or more underlying characteristics. See Appendix B for details on the reweighting process.

The DFL decomposition focuses on the contributions of Black-white differences in observed characteristics ("quantities"), rather than Black-white differences in returns to characteristics ("prices"). If Black men receive lower labor market returns (such as less working time and/or lower wage rates) to the same characteristics, this dimension of racial differences will be left in the residuals of DFL decomposition.[30]

In my empirical analysis, I present both the unconditional contributions of skills and the contributions of skills after conditioning on racial differences in family and/or neighborhood characteristics. In the DFL decomposition, estimating the conditional contributions of skills is intuitively equivalent to first residualizing the racial skill gap by racial differences in family and/or neighborhood characteristics and then estimating the contribution of the residualized racial skill gap.[31] If the racial skill gaps can be fully (or more than fully) explained by racial differences in family and/or neighborhood characteristics, the estimated conditional contribution of skills could be close to zero (or negative).[32]

When estimating the role of the school-to-work transition, I condition on racial differences in pre-market characteristics, because presumably the school-to-work transition is in part a function of these characteristics. For example, part of the Black-white gap in the school-to-work transition outcomes could be attributed to differences in education and skills accumulated before labor market entry. Similarly, Black-white differences in childhood neighborhood could affect residential choices in early adulthood, which could further affect Black-white differences in exposure to local unemployment rates upon labor market entry. Conditioning on pre-market characteristics before estimating the contributions of the school-to-work transition makes it more likely that the estimated contribution of the transition reflects the role of the transition itself, rather than reflecting racial differences in the pre-market characteristics.

[30]It might be of particular interest to further decompose the residuals to see, for example, the specific contribution of racial differences in skill prices. Doing this requires imposing additional structure and assumptions on the residuals and is beyond the scope of this paper. One example is Firpo, Fortin, and Lemieux (2018), who propose a decomposition method based on re-centered influence function regressions. Thompson (2021) uses linear regressions to examines the contributions of changing skill prices to changing racial gaps.

[31]Note that this is different from the classical Oaxaca-Blinder decomposition and from other decomposition methods based on linear regressions. In the Oaxaca-Blinder decomposition, when estimating the contributions of skills conditioning on family/neighborhood characteristics, different coefficients will be used (compared to when estimating the unconditional contributions of skills), but the raw racial skill gap (not the residualized racial skill gap) will still be used to compute the contributions of skills. See Appendix B for a detailed discussion.

[32]In practice, it is not uncommon for DFL decomposition to produce estimates with a negative sign (e.g., DiNardo, Fortin, and Lemieux, 1996; Altonji, Bharadwaj, and Lange, 2012).

5 Decomposition Results

The DFL decomposition shows how much of the racial employment and earnings gaps, measured at the sixth to eighth years, can be explained by racial differences in quantities of underlying characteristics. I perform the decomposition separately for the two NLSY cohorts and focus on the dual roles of education and skills and the school-to-work transition.[33] A large body of research has used the NLSY–79 data to examine the roles of different explanatory factors in understanding racial labor market gaps (e.g., Neal and Johnson, 1996; Altonji and Blank, 1999; Heckman, Stixrud, and Urzúa, 2006). I present the results for the NLSY–79 as a confirmation of this literature, and as a benchmark to understand what has and has not changed in the NLSY–97. I establish two main findings, which I detail in the following sections.

5.1 Role of Education and Skills

My first finding concerns the role of education and skills accumulated prior to labor market entry in explaining racial gaps in employment and earnings. I start by presenting results for the NLSY–97 cohort, as summarized in Table 3. In column (1), I first present the racial gap in employment (average weeks worked per year) and earnings (log of average annual earnings including years of zero earnings). In columns (2)–(4), I present the share of the gaps that can be explained by education and skills, under different specifications. In column (5), I present the share of the gaps explained by pre-market factors together (education and skills, family background, childhood and neighborhood), and in column (6) I present the share that is left unexplained and is in the residuals.

As column (2) shows, when estimated unconditionally, racial differences in education and skills explain 27% of the racial employment gap and 30% of the racial earnings gap in the NLSY–97. To understand whether this explanatory power of education and skills can be attributed

[33] A common practice in estimating wage or earnings equations is to control for work experience, which is (in many cases) approximated by age or potential experience. I do not include experience in my decomposition (or adjust labor market outcomes by experience) for three reasons. First, work experience itself is a potential outcome of the pre-market characteristics that my decomposition focuses on. Second, my measure of the school-to-work transition can be seen as a measure of labor market experience, but only in one's first year after schooling completion. Compared to total work experience, the transition measure is arguably more exogenous to the labor market outcomes measured over the sixth to eighth years post-schooling. Third, my NLSY samples include men who completed schooling for at least eight years (i.e., have potential experience for at least eight years). This already accounts for some of the racial differences in experience. If experience still matters in explaining racial labor market gaps, conditioning on pre-market characteristics and the school-to-work transition, its contribution will be captured in the residuals of my decomposition.

to racial differences at the family and neighborhood levels, I first estimate the contributions of education and skills after conditioning on family background (column 3) and again after conditioning on both family background and childhood neighborhood characteristics (column 4).

Results in columns (3) and (4) show that the explanatory power of education and skills in the younger cohort (NLSY–97) is quantitatively robust even after conditioning on measured racial differences at family and neighborhood levels. Conditioning on racial differences in family background, the explanatory power of education and skills is only slightly reduced, now explaining 25% of the racial employment gap and 29% of the racial earnings gap. Further conditioning on racial differences in childhood neighborhood characteristics, the explanatory power of education and skills barely changes, to 24% of the racial employment gap and 33% of the racial earnings gap.

I bootstrap the decomposition process 5,000 times to construct p-values and standard errors. In Appendix Table A.4, I show that the racial employment and earnings gaps explained by education and skills are statistically significant at the 5% level, both unconditionally and after conditioning on measured racial differences at the family and neighborhood levels. What does this stable explanatory power of education and skills suggest? I postpone answering this question to later when I draw a direct comparison to the NLSY–79 cohort.

Education and skills, family background, and childhood neighborhood together account for 52% of the racial employment gap and 61% of the racial earnings gap, as presented in column (4). Appendix Table A.3 presents the complete decomposition results, including estimated contributions of family and neighborhood characteristics. The racial employment and earnings gaps that are not explained by measured racial differences in the three pre-market characteristics are left in the residuals, as shown in column (5). Note that I leave the discussion of the school-to-work transition to the next section. In Tables 3–5, the contributions of the school-to-work transition are included in the residuals.

Given the quantitatively large and robust effect of education and skills in the NLSY–97, a natural question is whether a specific skill measure has driven this result. Recall that my set of skill measures includes highest grade completed, measured cognitive skills (AFQT score), and measured non-cognitive and social skills. Significant Black-white gaps are observed for most of the skill measures in the NLSY–97 (Table 2). Although it is difficult to disentangle the effects of different skill measures, since they can be endogenous to each other, I can explore, in a descriptive sense, which specific skill measure has the dominant explanatory power.[34]

Table 4 contrasts the decomposition results when only the highest grade completed is included in the set of education and skills (top panel) with the results where only cognitive skill (AFQT score) is included in the set of education and skills (bottom panel).[35] Each column of Table 4 presents the decomposition results in a similar structure as in Table 3. Comparing the top panel with the bottom panel, the explanatory power of education and skills is attributable primarily to measured cognitive skills (AFQT score) rather than to formal schooling (highest grade completed). The racial gap in highest grade completed unconditionally accounts for 12%–14% of the racial employment and earnings gaps, while its explanatory power decreases substantially to almost zero (1%–3%) after conditioning on racial differences in family background and childhood neighborhood characteristics.

In stark contrast, the racial gap in measured cognitive skills (AFQT score) unconditionally explains about 30% of the racial labor market gaps, and this explanatory power stays largely stable after conditioning on racial differences in family and neighborhood characteristics. The share of the racial labor market gaps accounted for by measured cognitive skills alone is very close to the share accounted for by the full set of skill measures, as previously shown in Table 3. The dominant explanatory power of AFQT score also highlights the necessity of including some appropriately constructed measure of cognitive skills when studying racial gaps, which is seldom available in "big data," such as administrative tax records.

After establishing the results for the younger cohort, I now show whether and how the role of education and skills has changed across cohorts. Table 5 replicates Table 3 in the top panel and contrasts it with the results of the NLSY–79 in the bottom panel. Neal and Johnson (1996) show that AFQT score (in a quadratic function) unconditionally accounts for about 60% of the wage gap between Black and white men in the NLSY–79. Consistent with their findings, column (2) in the bottom panel of Table 5 shows that racial differences in education and skills unconditionally explain 64% of the racial employment gap and 67% of the racial earnings gap in my NLSY–79 sample.

[34]In a structural model, Urzúa (2008) emphasizes the key insight that observed (AFQT) test scores are a function of both underlying (cognitive) ability and other characteristics, including family background characteristics (such as parental income). The paper shows that cognitive skills can grow as education attainment increases and people can make endogenous schooling decisions based on their underlying cognitive and non-cognitive abilities.

[35]The results shown here barely change when I add non-cognitive and/or social scores.

There are two major changes between the NLSY cohorts. First, the unconditional explanatory power of education and skills is higher in the older cohort (more than 60%) than in the younger cohort (about 30%). As discussed earlier in Table 2, the racial gap in measured cognitive skills (AFQT score) fell substantially and statistically significantly from 35.7 percentiles in the NLSY–79 cohort to 25.3 percentiles in the NLSY–97 cohort. If the returns to cognitive skills stayed stable across cohorts, a smaller racial gap in cognitive skills would lead to a lower explanatory power of education and skills in the younger cohort.[36] In contrast, recall that in Table 2, the racial gap increased insignificantly for highest grade completed and significantly for social score. Because the explanatory power of education and skills is primarily driven by AFQT score, rather than by highest grade completed or social score, the falling racial gap in AFQT score is playing the dominant role.[37]

Second, the explanatory power of education and skills in the NLSY–79 decreases substantially after accounting for racial differences at family and neighborhood levels, while the explanatory power of education and skills in the NLSY–97, as previously discussed, is broadly stable before and after conditioning on family and neighborhood characteristics. When estimated conditioning on racial differences in family background (column 3 in the bottom panel of Table 5), the share explained by education and skills in the NLSY–79 decreases by more than half (from 64% to 30%) for the racial employment gap and decreases by more than two-thirds (from 67% to 17%) for the racial earnings gap. When estimated conditioning on both family and neighborhood characteristics (column 4 in the bottom panel of Table 5), the share of the racial labor market gaps explained by education and skills in the NLSY–79 further drops to below zero.[38]

For the NLSY–79 cohort, bootstrap results in Appendix Table A.7 show that the explanatory power of education and skills is significant at the 5% level when estimated unconditionally, but

[36]Note that the DFL method abstracts from estimating the specific relationships between the outcomes and the explanatory factors. That is, the method does not directly estimate returns to skills or to other factors.

[37]If labor market returns to education and skills were to decrease across cohorts, the explanatory power of education and skills would also decrease (assuming racial differences in education and skills have remained unchanged). However, existing evidence on how returns to education and skills have changed in the U.S. labor market is mixed and still preliminary. Castex and Dechter (2014) find that the returns to education have increased and the returns to cognitive skills have decreased from the NLSY–79 cohort to the NLSY–97 cohort. Deming (2017) further shows that the returns to non-cognitive skills and social skills have increased between the two cohorts. That said, Hellerstein, Luo, and Urzúa (2019) show that the former two studies rely on a strong assumption of constant skill prices, an assumption that does not seem to hold in the NLSY–97 cohort. When relaxing this assumption, there is no conclusive evidence on a cross-cohort decline in the returns to cognitive skills.

[38]As discussed in Section 4, the estimated "negative" contributions of education and skills here suggest that the racial skill gaps in the NLSY–79 can be more than fully compensated by the combined racial differences in family and neighborhood characteristics.

becomes indistinguishable from zero after conditioning on racial differences in family background and childhood neighborhood characteristics.[39]

The decomposition results of the NLSY–79 suggest that much (if not all) of the measured racial skill gaps in the older cohort can be explained by racial differences at family and neighborhood levels. Past studies have discussed the roles of family, neighborhood, and other factors in shaping the skill development process.[40] For example, using a linear regression, Neal and Johnson (1996) show that young men in the NLSY–79 who have high AFQT scores are from a more advantageous background (e.g., more highly educated parents, reading materials at home) and a better school environment (e.g., lower student-to-teacher ratio, lower student dropout rate).

However, this pattern changes in the NLSY–97, and the racial skill gaps in the younger cohort do not seem to have originated from racial differences in family and neighborhood characteristics, at least as far can be measured in the NLSY data. This leads to a further question of where the racial skill gap in the younger cohort originates.

To understand the origins of the observed racial skill gaps, it is important to emphasize that the skill measures themselves can be seen as an outcome. For example, the cognitive skills in my data (AFQT score) are measured when respondents were ages 12–18 and could be a function of a series of family investments, school influences, and/or neighborhood impacts that happened in early childhood years.[41] Identifying the specific mechanisms behind the racial skill differences in the younger cohort is beyond the scope of this paper, but existing studies can shed light on what the potential mechanisms may and may not be.

In a cohort close in age to the NLSY–97, Chetty et al. (2020) show descriptively that low-poverty neighborhoods (census tracts) with low levels of racial bias among whites and high rates of father presence among Blacks tend to have smaller racial income gaps. Considering the

[39] The complete decomposition results for the NLSY–79, including the estimated contributions of family and neighborhood characteristics, are presented in Appendix Table A.6.

[40] See Cunha et al. (2006) for a review.

[41] The AFQT score is measured at ages 15–18 in my final sample of the NLSY–79 cohort and ages 12–18 in the NLSY–97 cohort. Note that I use the score constructed by Altonji, Bharadwaj, and Lange (2012), which carefully concords the two cohorts to make the AFQT scores comparable. According to Deming (2017), the non-cognitive score is constructed from two tests (one conducted at ages 14–17 and one at ages 15–18) in the NLSY–79 cohort and is constructed from two sets of questions (one asked at ages 17–21 and one at ages 23–27) in the NLSY–97 cohort. The social score is constructed from two sets of questions (one aims to measure sociability in high school, and one aims to measure sociability at age 6 and as an adult) in the NLSY–79 cohort and is constructed from a set of questions asked at ages 23–27 in the NLSY–97 cohort. As discussed earlier, there is no evidence that the non-cognitive and social scores are directly comparable between the two NLSY cohorts. The findings in this paper are quantitatively robust when excluding non-cognitive and social scores from the analysis.

relative role of families and neighborhoods (including schools) in the skill formation of children, past studies have established that family investments play a much more crucial role than school and neighborhood influences (Cunha et al., 2006).[42]

As a brief investigation into this, in results not shown here, I include a richer set of family and neighborhood variables available only in the NLSY–97 data. The added family variables include mother parenting style and teenage mom status, and the added neighborhood variables include a measure of county neighborhood quality (Chetty and Hendren, 2018b) and homeownership status. After further conditioning on racial differences in these family and neighborhood variables, the estimated contributions of education and skills are still quantitatively robust in the NLSY–97. A more comprehensive investigation of the skill accumulation process for the younger cohort requires future research.

5.2 Role of the School-to-Work Transition

My second finding focuses on the role of the school-to-work transition in explaining racial gaps in future labor market outcomes, on top of the roles of pre-market skills and family and neighborhood characteristics. As shown earlier in Table 1, Black men in both cohorts had worse outcomes than their white counterparts in various labor market dimensions in their very first year post-schooling. It has been widely documented that the school-to-work transition has a long-lasting impact on future labor market outcomes (Neumark, 2002; Kahn, 2010), especially for minority and economically disadvantaged groups (Schwandt and Wachter, 2019).[43] If Black disadvantage caused by the school-to-work transition persists through early career years, it may help explain some of the racial labor market gaps observed in the sixth to eighth years post-schooling, even after conditioning on racial differences in the pre-market characteristics.

Figures 1–2 and Table 1 reveal some suggestive evidence on the persistent effect of the school-to-work transition. In the NLSY–79 cohort, the Black-white gap in the transition stage (defined as the first year post-schooling) shows some convergence over the following four to five years. In particular, the convergence is statistically significant for weeks worked per year. In the NLSY–97

[42]On top of the roles that family, school, and neighborhood play, pervasive racial discrimination against Black men can also affect the skill formation process. For example, Black families and children anticipating future discrimination in the labor market may choose to underinvest in education and skills.

[43]Rinz (2019) shows that exposure to the Great Recession has cost Black workers 1.33 years of their average earnings and has cost white workers 0.94 years of their average earnings. But the estimates are based on all workers, not new workers who just entered the labor market.

cohort, there was much less convergence, and the initial racial gaps largely persisted over the early career years. Although this is not a causal estimate, it does suggest that the persistent impact of the school-to-work transition is especially relevant for the NLSY–97 cohort.

In the decomposition, I employ two different measures of the school-to-work transition. The first measure is a flexible function of weeks worked in the first year post-schooling. The second measure is local unemployment rates at one's place of residence upon school completion. County-level unemployment rate data are unavailable for most of the NLSY–79 sample years, and I only use them for the NLSY–97. Again, I condition my estimates on racial differences in education and skills, family background, and childhood neighborhood characteristics.

Table 6 presents the estimated explanatory power of the school-to-work transition separately for the NLSY–97 (top panel) and the NLSY–79 (bottom panel). Within each panel, column (1) shows the racial gaps in employment and earnings, column (2) shows the share of the gap explained by the pre-market factors (education and skills, family background, childhood neighborhood) together, and column (3) shows the share of the gap that can be explained by the racial differences in weeks worked in the first year post-schooling, conditioning on racial differences in the pre-market characteristics. Column (4) further adds unemployment rates at entry state-year to my set of transition measures (on top of weeks worked in the first year post-schooling), and presents the share of the racial gaps that can be explained in addition to the share explained in column (3). Similarly, column (5) adds unemployment rates at entry county-year, and presents the share of the racial gaps that can be explained in addition to the share explained in column (3).

In the younger cohort, when the school-to-work transition is measured by weeks worked in the first year post-schooling (top panel column 3), the transition accounts for about 13% of the racial employment gap and 13% of the racial earnings gap, after conditioning on racial differences in pre-market characteristics. When the school-to-work transition is measured by unemployment rates at entry county-year *together with* weeks worked in the first year (top panel column 5), it accounts for 31% (13% + 18% = 31%) of the racial employment gap and 34% (13% + 21% = 34%) of the racial earnings gap, after conditioning on the pre-market characteristics. Out of this explanatory power of transition, more than half comes from racial differences in the exposure to different county unemployment rates upon schooling completion.[44]

[44] In results not presented here, when unemployment rates at entry county-year are added alone as a measure of transition (i.e., excluding weeks worked in the first year post-schooling), this accounts for about 20% of the

Using state unemployment rates upon schooling completion (top panel column 4), which is a less accurate measure of local labor market conditions, does not achieve explanatory power close to that of county unemployment rates.

In stark contrast, in the older cohort (NLSY–79), after conditioning on racial differences in pre-market characteristics, accounting for racial differences in transition adds no extra explanatory power, whether I measure the transition by weeks worked in the first year post-schooling (bottom panel column 3) or by the state unemployment rate upon schooling completion (bottom panel column 4).[45]

Why is the role of the school-to-work transition substantially more important in the younger cohort? Distinct from previous cohorts of Americans, the NLSY–97 cohort (early millennials) went through their early career years under the long-lasting shadow of the Great Recession. As shown in Appendix Figure A.1, most of the Black and white men in my NLSY–97 sample spent at least part of their first eight post-schooling years between 2008 and 2015. Although the NLSY–79 cohort also experienced smaller recessions over their early careers, it is possible that the Great Recession was particularly destructive for the job prospects of young men, and failing to transition smoothly from school to work was especially costly during the Great Recession.

In addition, it is important to emphasize that Black men completed schooling (and started their careers) at a location and time with worse labor market conditions, for complicated reasons that can extend beyond simply bad luck. Past research has shown that Black workers tend to live in places with fewer job opportunities, and the relocation of firms from central cities to suburban rings has paralleled the declining Black employment rate in central cities (Hellerstein and Neumark, 2012; Miller, 2018). The observed Black disadvantage in the school-to-work transition could be, at least partially, due to barriers to geographic mobility, a lack of resources to freely choose their school-leaving time, and eventually a lack of access to job opportunities. It could also be due to discrimination against Black men in the hiring process and at the workplace, which further discourages Black men from searching and/or migrating for work.

racial labor market gaps in the NLSY–97, conditioning on the pre-market factors.

[45] In the NLSY–79, as shown in Table 6, the explanatory power of transition conditioning on pre-market factors is negative in the DFL decomposition. As explained in Section 4 and Appendix B, this is because racial differences in the pre-market characteristics *more than fully* account for racial differences in the school-to-work transition as measured in the NLSY–79.

6 Conclusion

How have the Black-white labor market gaps among young men changed across cohorts, and how have the underlying forces of these gaps changed? Given that both the characteristics of Americans and the overall structure of the labor market have changed dramatically in the past several decades (Altonji, Bharadwaj, and Lange, 2012; Autor and Dorn, 2013; Castex and Dechter, 2014; Deming, 2017), one cannot simply assume that our knowledge based on previous cohorts applies to today's cohort of young men. In this paper, I provide some of the first evidence on these questions with the help of two similarly constructed and nationally representative samples of young Americans: the NLSY–97 and the NLSY–79.

Tracking the early career trajectories of Black and white men in the two cohorts, I first find that the upward-sloping employment and earnings trajectories observed in the older cohort (NLSY–79) is dampened in the younger cohort (NLSY–97), especially for employment outcomes of Black men. In the older cohort, the racial employment gap in the transition stage (measured by weeks worked in the first year after completing school) narrows substantially and significantly over the first six to eight years post-schooling. But in the younger cohort, this narrowing of the initial racial employment gap is quantitatively much smaller and statistically insignificant over the early career years.

I then explore whether and how the explanatory factors underlying the observed racial labor market gaps have changed between the two NLSY cohorts, focusing on the roles of education and skills and the school-to-work transition. First, education and skills, especially measured cognitive skills, explain a crucial part (30%) of the racial labor market gaps in the younger cohort (NLSY–97). Although the explanatory power of skills decreases across cohorts as the racial skill gap narrows, the explanatory power of education and skills is robust in the younger cohort, even after accounting for measured racial differences in family and neighborhood characteristics.

Second, racial differences in the school-to-work transition (measured by weeks worked in the first year post-schooling and local labor market conditions upon schooling completion) play a particularly important role in explaining racial labor market gaps in the younger cohort (NLSY–97). This is possibly driven by the fact that the younger cohort spent much of their early career years under the shadow of the Great Recession, which made an unsuccessful school-to-work transition especially consequential.

Given the descriptive nature of my findings, one must be cautious in drawing immediate policy implications. However, combining my findings with existing studies suggests lessons that may help guide future research and polices. Despite the dramatic changes in both the characteristics of young men and the overall structure of the U.S. labor market, cognitive skills turn out to still be the key driver of racial labor market gaps in the younger cohort, as in previous cohorts. This finding suggests that, although market demand for skills might have evolved over the past few decades, cognitive skills are still rewarded in today's labor market and are particularly important in shaping racial gaps in labor market outcomes.

Although the racial gap in cognitive skills (measured by AFQT score) has narrowed across the two NLSY cohorts, the gap remains quantitatively substantial and statistically significant in the younger cohort (NLSY–97). My findings strongly suggest that more attention needs to be paid the skill accumulation process and, more importantly, institutional and economic barriers to Black men and families in this process. Potentially effective pathways to reduce racial labor market gaps include public programs that foster skill accumulation among Black men, especially before labor market entry.[46]

My finding regarding the role of the school-to-work transition suggests that helping disadvantaged Black men get a foothold in the labor market is a potentially important pathway to reduce racial labor market gaps at later career stages. In addition to traditional government training programs[47], recent examples of job training programs designed and led by non-government organizations show encouraging results for helping disadvantaged youths initiate a career (Fein and Hamadyk, 2018).[48] Policies that increase geographic mobility and the flexibility to choose school-leaving time may also help reduce Black disadvantage in the school-to-work transition process, which will eventually help reduce racial gaps in longer-term labor market outcomes.

References

Altonji, Joseph, Prashant Bharadwaj, and Fabian Lange (2012). "Changes in the Characteristics of American Youth: Implications for Adult Outcomes". In: *Journal of Labor Economics.*

[46]For example, evidence based on previous cohorts suggests that family investments in young children have especially high returns (Cunha et al., 2006).

[47]Friedlander, Greenberg, and Robins (1997) reviews the evidence on the effectiveness of government training programs and concludes that the effects are modest on average and limited among youths.

[48]I thank Peter Bergman for pointing out the innovative job training programs (e.g., Year Up) led by non-government organizations.

Altonji, Joseph and Rebecca Blank (1999). “Race and gender in the labor market”. In: *Handbook of Labor Economics*. Elsevier.

Autor, David and David Dorn (2013). “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market”. In: *American Economic Review*.

Barsky, Robert et al. (2002). “Accounting for the Black-White Wealth Gap: A Nonparametric Approach”. In: *Journal of the American Statistical Association*.

Bayer, Patrick and Kerwin Charles (2018). “Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940”. In: *The Quarterly Journal of Economics*.

Bertrand, Marianne and Sendhil Mullainathan (2004). “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination”. In: *American Economic Review*.

Blinder, Alan (1973). “Wage Discrimination: Reduced Form and Structural Estimates”. In: *Journal of Human Resources*.

Castex, Gonzalo and Kogan Dechter (2014). “The Changing Roles of Education and Ability in Wage Determination”. In: *Journal of Labor Economics*.

Charles, Kerwin and Jonathan Guryan (2008). “Prejudice and Wages: An Empirical Assessment of Becker’s The Economics of Discrimination”. In: *Journal of Political Economy*.

Chetty, Raj and Nathaniel Hendren (2018a). “Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects”. In: *The Quarterly Journal of Economics*.

— (2018b). “Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates”. In: *The Quarterly Journal of Economics*.

Chetty, Raj, Nathaniel Hendren, Maggie Jones, et al. (2020). “Race and Economic Opportunity in the United States: An Intergenerational Perspective”. In: *The Quarterly Journal of Economics*.

Chetty, Raj, Nathaniel Hendren, and Lawrence Katz (2016). “The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment”. In: *American Economic Review*.

Chyn, Eric (2018). “Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children”. In: *American Economic Review*.

Council of Economic Advisors (2016). *Economic Perspectives on Incarceration and the Criminal Justice System*. Tech. rep.

Cunha, Flavio et al. (2006). "Interpreting the Evidence on Life Cycle Skill Formation". In: *Handbook of Economics of Education*. Elsevier.

Daly, Mary, Bart Hobijn, and Joseph Pedtke (2017). "Disappointing Facts about the Black-White Wage Gap". In: *Federal Reserve Bank of San Francisco Economic Letter*.

Deming, David (2017). "The Growing Importance of Social Skills in the Labor Market". In: *The Quarterly Journal of Economics*.

DiNardo, John, Nicole Fortin, and Thomas Lemieux (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach". In: *Econometrica*.

Donohue, John and James Heckman (1991). "Continuous Versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks". In: *Journal of Economic Literature*.

Fein, David and Jill Hamadyk (2018). *Bridging the Opportunity Divide for Low-Income Youth: Implementation and Early Impacts of the Year Up Program*. Tech. rep.

Firpo, Sergio, Nicole Fortin, and Thomas Lemieux (2018). "Decomposing Wage Distributions Using Recentered Influence Function Regressions". In: *Econometrics*.

Forbes (2016). "Millennials Are Struggling To Get Jobs". In: URL: `https://www.forbes.com/sites/larryalton/2016/12/22/millennials-are-struggling-to-get-jobs-heres-why-and-what-to-do-about-it/`.

Friedlander, Daniel, David Greenberg, and Philip Robins (1997). "Evaluating government training programs for the economically disadvantaged". In: *Journal of Economic Literature*.

García, Jorge Luis et al. (2020). "Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program". In: *Journal of Political Economy*.

Gelbach, Jonah (2016). "When Do Covariates Matter? And Which Ones, and How Much?" In: *Journal of Labor Economics*.

Heckman, James (1998). "Detecting Discrimination". In: *Journal of Economic Perspectives*.

— (2018). *Discussion on Raj Chetty's Lecture at the Princeton University*. `https://uchv.princeton.edu/events/raj-chetty-stanford-university-lecture-ii-neighborhood-effects-childhood-environment-and`. Accessed: 2019-10-01.

Heckman, James, Jora Stixrud, and Sergio Urzúa (2006). "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior". In: *Journal of Labor Economics*.

Hellerstein, Judith, Sai Luo, and Sergio Urzúa (2019). "Have the Returns to Cognitive Skills Really Declined?" In: *Working paper*.

Hellerstein, Judith and David Neumark (2012). "Employment in Black Urban Labor Markets: Problems and Solutions". In: *The Oxford Handbook of the Economics of Poverty*. National Bureau of Economic Research.

Heywood, John and Daniel Parent (2012). "Performance Pay and the White-Black Wage Gap". In: *Journal of Labor Economics*.

Kahn, Lisa (2010). "The Long-term Labor Market Consequences of Graduating from College in a Bad Economy". In: *Labour Economics*.

Lang, Kevin and Michael Manove (2011). "Education and Labor Market Discrimination". In: *American Economic Review*.

Light, Audrey and Kathleen McGarry (1998). "Job Change Patterns and the Wages of Young Men". In: *Review of Economics and Statistics*.

Miller, Conrad (2018). "When Work Moves: Job Suburbanization and Black Employment". In: *National Bureau of Economic Research Working Paper Series*.

Neal, Derek and William Johnson (1996). "The role of premarket factors in black-white wage differences". In: *Journal of Political Economy*.

Neal, Derek and Armin Rick (2014). "The Prison Boom and the Lack of Black Progress after Smith and Welch". In: *NBER Working Paper*.

Neumark, David (2002). "Youth Labor Markets in the United States: Shopping Around vs. Staying Put". In: *Review of Economics and Statistics*.

Oaxaca, Ronald (1973). "Male-Female Wage Differentials in Urban Labor Markets". In: *International Economic Review*.

Pager, Devah (2003). "The Mark of a Criminal Record". In: *American Journal of Sociology*.

Prada, Marıa F and Sergio Urzúa (2017). "One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance, and Earnings". In: *Journal of Labor Economics*.

Rinz, Kevin (2019). "Did Timing Matter? Life Cycle Differences in Effects of Exposure to the Great Recession". In: *Working Paper*.

Rodgers, William and William Spriggs (1996). “What does the AFQT really measure: Race, wages, schooling and the AFQT score”. In: *The Review of Black Political Economy.*

Schwandt, Hannes and Till von Wachter (2019). “Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets”. In: *Journal of Labor Economics.*

Smith, James and Finis Welch (1989). “Black Economic Progress After Myrdal”. In: *Journal of Economic Literature.*

The Atlantic (2015). “Even Baby Boomers Think It’s Harder to Get Started Than It Used to Be”. In: URL: https://www.theatlantic.com/business/archive/2015/06/even-baby-boomers-think-its-harder-to-get-started-than-it-used-to-be/395609/.

Thompson, Owen (2021). “Human Capital and Black-White Earnings Gaps, 1966-2017”. In: *Working paper.*

Topel, Robert and Michael Ward (1992). “Job Mobility and the Careers of Young Men”. In: *The Quarterly Journal of Economics.*

Urzúa, Sergio (2008). “Racial Labor Market Gaps: The Role of Abilities and Schooling Choices”. In: *Journal of Human Resources.*

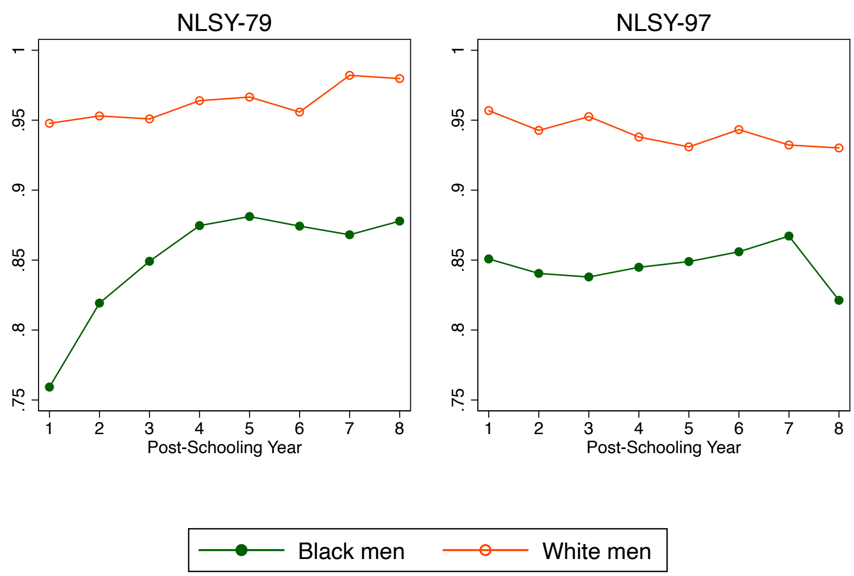
Wigdor, Alexandra and Bert Green Jr. (1991). *Performance Assessment for the Workplace.* National Academy Press.

Wilson, Valerie and William Rodgers (2016). *Black-white wage gaps expand with rising wage inequality.* Tech. rep.

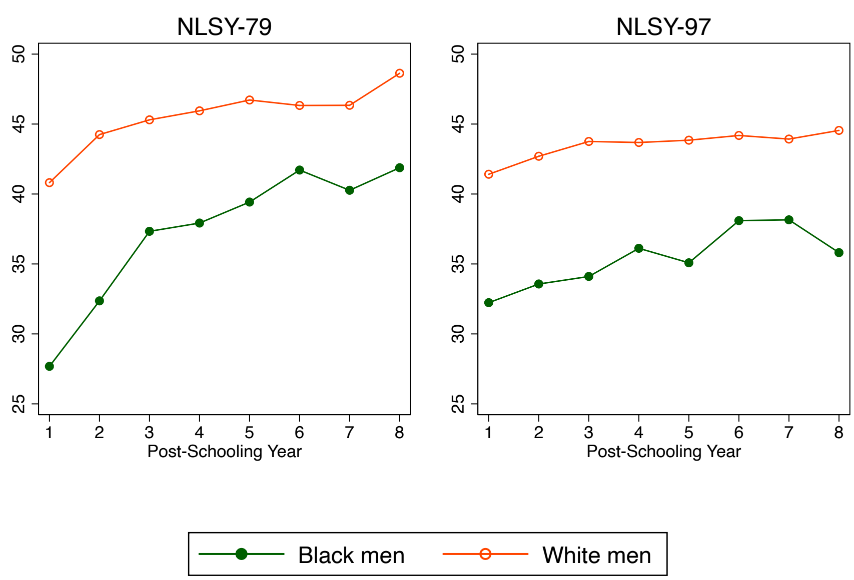
Figures and Tables

Figure 1: Career Trajectories: Any Employment and Weeks Worked

(a) Any Employment



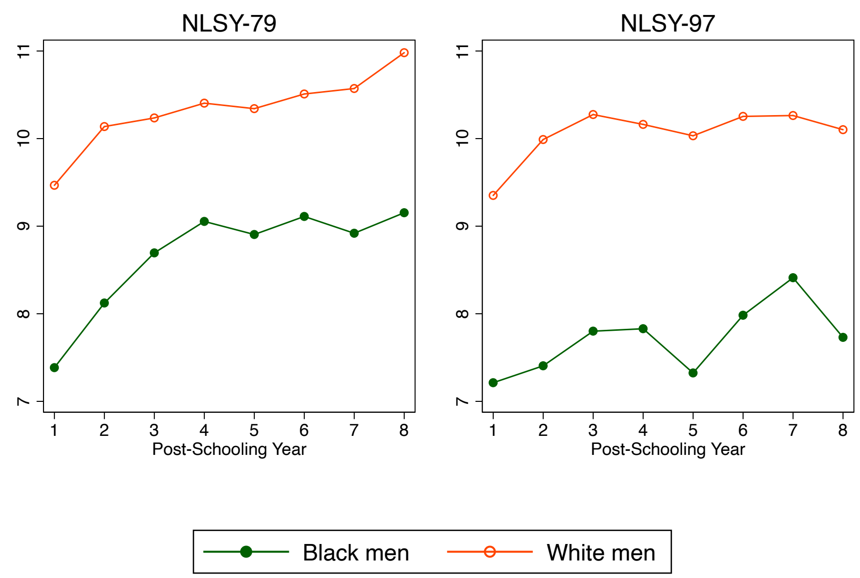
(b) Weeks Worked



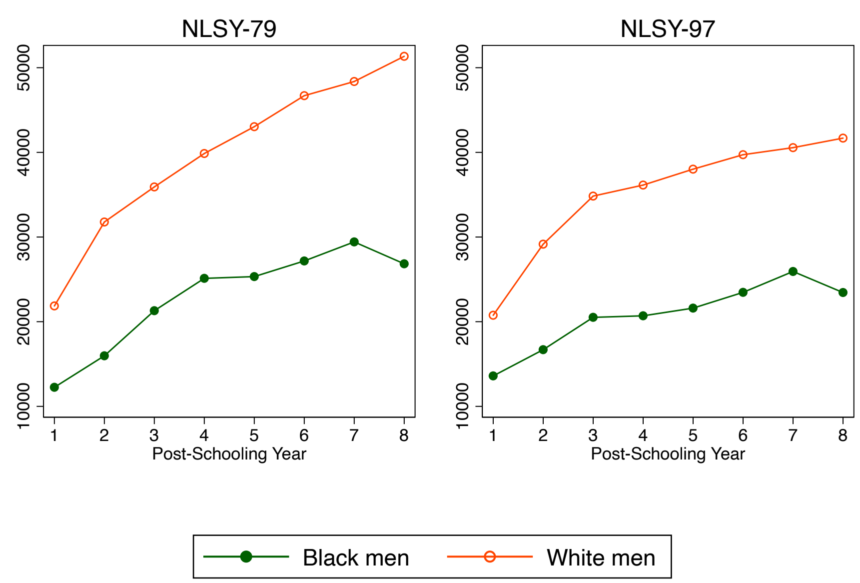
Notes: Both NLSY–79 and NLSY–97 samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used.

Figure 2: Career Trajectories: Annual Earnings

(a) Log Annual Earnings



(b) Annual Earnings



Notes: Annual earnings are adjusted to 2013 dollars. The top panel takes the inverse hyperbolic sine of annual earnings. Both NLSY–79 and NLSY–97 samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used.

Table 1: Early Career Outcomes of Black and White Men in the NLSY–79 and NLSY–97 Cohorts

	NLSY–79			NLSY–97			97–79
	White	Black	W-B	White	Black	W-B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Transition Stage (1st Year)							
Age at school completion	20.44	19.88	0.56†	19.54	19.28	0.26	-0.30
Weeks before finding 1st job	11.89	41.97	–30.08†	10.68	32.78	–22.10†	7.98
Any employment	0.95	0.76	0.19†	0.96	0.85	0.11†	–0.08†
Worked for $\geq$ 26 weeks	0.80	0.54	0.26†	0.85	0.65	0.20†	–0.06
Worked for $\geq$ 50 weeks	0.57	0.29	0.28†	0.48	0.35	0.13†	–0.15†
Weeks worked	40.81	27.68	13.13†	41.41	32.23	9.18†	–3.95
Log annual earnings	9.47	7.38	2.08†	9.35	7.21	2.14†	0.06
Later Stage (6th–8th Year)							
Weeks worked per year	47.10	41.33	5.78†	44.20	37.30	6.90†	1.12
Log average annual earnings	10.93	9.56	1.37†	10.55	9.01	1.54†	0.17
Growth from 1st to 6th–8th Year							
Weeks worked per year	6.29	13.65	-7.36†	2.80	5.07	-2.27	5.09†
Log average annual earnings	1.48	2.14	-0.66	1.20	1.79	-0.58	0.58
Summarizing First 8 Years							
Number of NE spells	1.68	2.30	–0.62†	1.77	2.66	–0.89†	–0.28
Avg. months of NE spells	6.99	9.51	–2.52	6.47	10.14	–3.66†	–1.14
Cumulative weeks worked	363.92	297.75	66.17†	347.15	282.56	64.59†	–1.58
Weeks worked per year	45.54	37.29	8.25†	43.50	35.40	8.10†	–0.15
Log average annual earnings	11.04	10.01	1.03†	10.73	9.48	1.25†	0.22

[1] The NLSY–79 sample includes 444 white men and 271 Black men, and the NLSY–97 sample includes 825 white men and 396 Black men. Both samples are balanced panels of men who have completed formal schooling for at least eight years. Section 2 explains how the samples are constructed. Sample weights are used.
[2] † indicates a p-value below 0.05.
[3] NE stands for non-employment.

Table 2: Descriptive Characteristics of Black and White Men in the NLSY–79 and NLSY–97 Cohorts

	NLSY–79			NLSY–97			97–79
	White (1)	Black (2)	W-B (3)	White (4)	Black (5)	W-B (6)	(7)
Education and Skills							
HGC	13.46	12.80	0.66†	13.13	12.08	1.05†	0.39
AFQT percentile	58.61	22.93	35.68†	51.06	25.76	25.31†	–10.37†
Social score	0.03	0.20	–0.18	–0.05	–0.22	0.17†	0.35†
Non-cognitive score	0.07	–0.06	0.13	–0.12	–0.07	–0.05	–0.18
Family Background							
Log parental income	11.56	10.85	0.72†	10.96	8.94	2.01†	1.30†
Mother's HGC	12.07	11.13	0.94†	13.03	12.45	0.59†	–0.35
Living with both parents	0.85	0.56	0.28†	0.62	0.32	0.31†	0.02
Childhood Neighborhood							
Residence Type							
MSA, central city	0.06	0.36	–0.30†	0.21	0.36	–0.15†	0.14†
MSA, non-central city, urban	0.58	0.39	0.20†	0.33	0.22	0.12†	–0.08
MSA, non-central city, rural	0.04	0.00	0.04†	0.20	0.19	0.01	–0.03
Non-MSA, rural	0.19	0.17	0.02	0.18	0.15	0.03	0.01
County Conditions							
Log population	12.26	12.60	–0.35†	12.14	12.35	–0.21	0.13
Log median HH income	10.93	10.78	0.15†	10.95	10.87	0.08†	–0.07†
Poverty rate	0.11	0.17	–0.06†	0.11	0.16	–0.04†	0.02†
Male college rate	0.19	0.18	0.01	0.24	0.23	0.01	0.00
State Conditions							
Log population	15.75	15.65	0.10	15.79	15.76	0.03	–0.07
Log median HH income	10.91	10.85	0.07†	10.98	10.94	0.04†	–0.03
Poverty rate	0.12	0.14	–0.03†	0.12	0.13	–0.01†	0.02†
Male college rate	0.20	0.19	0.01†	0.26	0.25	0.01†	0.00

[1] HGC stands for highest grade completed. AFQT stands for the Armed Forces Qualification Test. MSA stands for metropolitan statistical area. HH stands for household. The NLSY–79 sample includes 444 white men and 271 Black men, and the NLSY–97 sample includes 825 white men and 396 Black men. Both samples are balanced panels of men who have completed formal schooling for at least eight years. Section 2 explains how the samples are constructed. Sample weights are used.

[2] † indicates a p-value below 0.05.

Table 3: DFL Decomposition (NLSY–97)

	W-B Gap	Share Explained by Skills			Pre-market	Residuals
		Uncond.	Cond. on Family	Cond. on Family & NBHD	Factors	
	(1)	(2)	(3)	(4)	(5)	(6)
NLSY–97						
Weeks worked per year	6.90	27%	25%	24%	52%	48%
Log avg annual earnings	1.54	30%	29%	33%	61%	39%

[1] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and NBHD stands for neighborhood. The sample is a balanced panel of 825 white men and 396 Black men who have completed formal schooling for at least eight years. Sample weights are used.
[2] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take the inverse hyperbolic sine.

Table 4: Is it Schooling or Measured Cognitive Skills? (NLSY–97)

	W-B Gap	Share Explained by Skills			Pre-market	Residuals
		Uncond.	Cond. on Family	Cond. on Family & NBHD	Factors	
	(1)	(2)	(3)	(4)	(5)	(6)
Skills include only HGC						
Weeks worked per year	6.90	14%	3%	1%	30%	70%
Log avg annual earnings	1.54	12%	3%	3%	31%	69%
Skills include only AFQT						
Weeks worked per year	6.90	28%	25%	23%	51%	49%
Log avg annual earnings	1.54	30%	26%	30%	59%	41%

[1] The top panel includes only highest grade completed (HGC) in the skill set, the bottom panel includes only Armed Forces Qualification Test (AFQT) score in the skill set. I use the AFQT score adjusted by Altonji, Bharadwaj, and Lange (2012).
[2] NBHD stands for neighborhood. The sample is a balanced panel of 825 white men and 396 Black men who have completed formal schooling for at least eight years. Sample weights are used.
[3] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take the inverse hyperbolic sine.

Table 5: Comparing DFL Decomposition of NLSY–97 with NLSY–79

	W-B Gap	Share Explained by Skills			Pre-market	Residuals
		Uncond.	Cond. on Family	Cond. on Family & NBHD	Factors	
	(1)	(2)	(3)	(4)	(5)	(6)
NLSY–97						
Weeks worked per year	6.90	27%	25%	24%	52%	48%
Log avg annual earnings	1.54	30%	29%	33%	61%	39%
NLSY–79						
Weeks worked per year	5.78	64%	30%	-37%	73%	27%
Log avg annual earnings	1.37	67%	17%	-25%	59%	41%

[1] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and NBHD stands for neighborhood. The NLSY–97 sample (top panel) includes 825 white men and 396 Black men, and the NLSY–79 sample (bottom panel) includes 444 white men and 271 Black men. Both samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used.

[2] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take the inverse hyperbolic sine.

Table 6: Role of the School-to-Work Transition

	W-B Gap	Pre-Market Factors	Share Explained by Transition Cond. on Pre-Market Factors		
			Weeks worked in 1st Year	Col (3) +State UR	Col (3) +County UR
	(1)	(2)	(3)	(4)	(5)
NLSY–97					
Weeks worked per year	6.90	52%	13%	3%	18%
Log avg annual earnings	1.54	61%	13%	3%	21%
NLSY–79					
Weeks worked per year	5.78	73%	–5%	–14%	
Log avg annual earnings	1.37	59%	–14%	–4%	

[1] UR stands for unemployment rates. County-level UR data, from the Local Area Unemployment Statistics (LAUS) program, go back to 1990 and is unavailable for most of my NLSY–79 sample years. The top panel uses the NLSY–97 sample and the bottom panel uses the NLSY–79 sample. Sample weights are used.

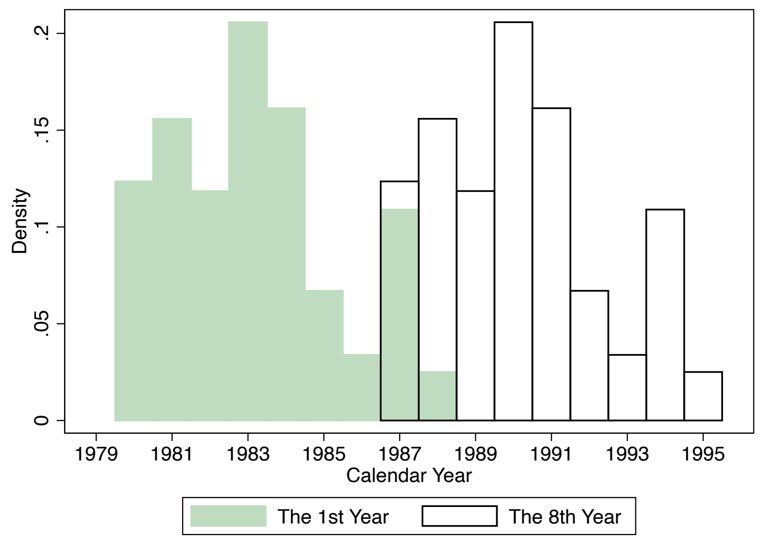
[2] Column 2 presents the overall explanatory power of pre-market characteristics (Skill, Family, Neighborhood). Columns 3-5 present the explanatory power of various school-to-work transition measures, estimated conditioning on the pre-market characteristics. Column 3 measures transition with a flexible vector of weeks worked in the first year post-schooling. Column 4, in addition to the transition measure in Column 3, adds UR in one's state of residence at the labor market entry year. Column 5, in addition to Column 3, adds UR in one's county of residence at the labor market entry year. The numbers shown in Columns 4-5 are the *additional* explanatory power of the specific transition measure, upon the explanatory power of weeks worked in the first year.

[3] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take inverse hyperbolic sine.

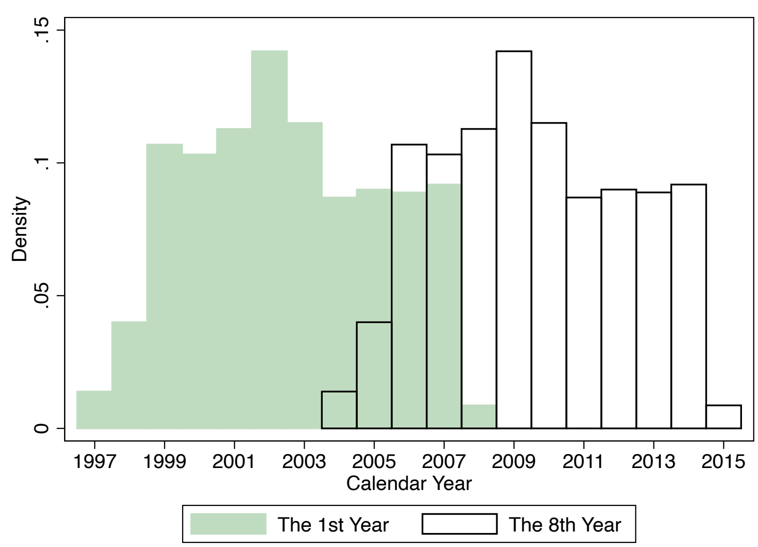
A Appendix Figures and Tables

Figure A.1: Corresponding Calendar Years to Sample Observations

(a) NLSY–79



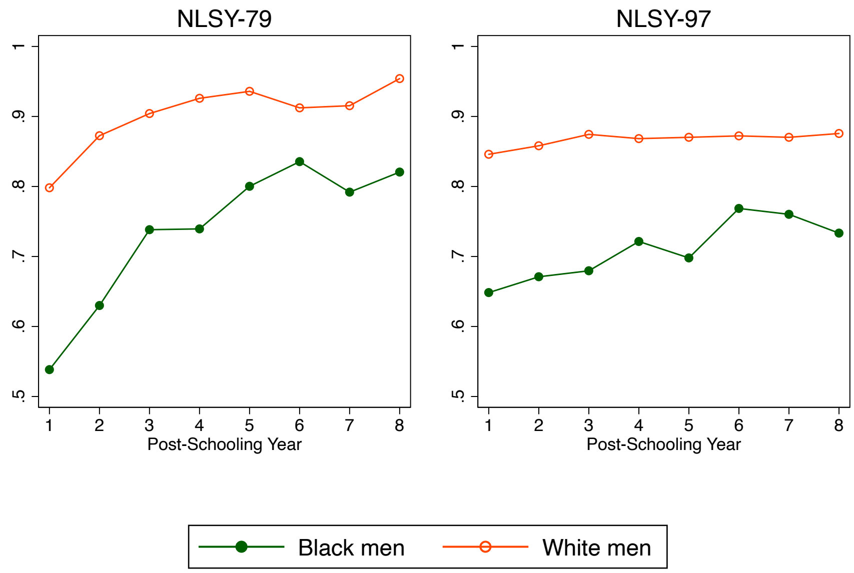
(b) NLSY–97



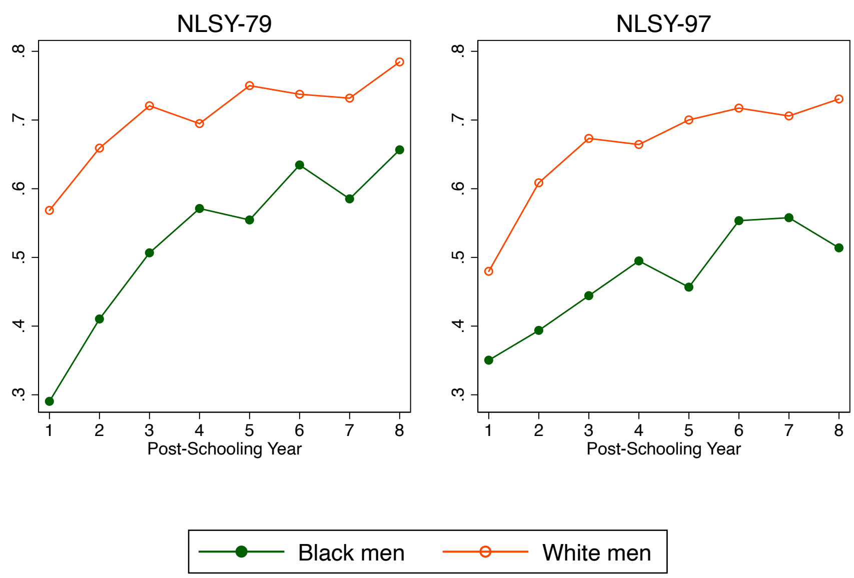
Notes: The histograms show the corresponding calendar years to the young men's first and eighth year post-schooling. Both NLSY–79 and NLSY–97 samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used.

Figure A.2: Career Trajectories: Worked Half Year and Full Year

(a) Worked Half Year ($\geq$ 26 weeks)



(b) Worked Full Year ($\geq$ 50 weeks)



Notes: Both NLSY–79 and NLSY–97 samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used.

Table A.1: Sample Construction and Sample Size

	NLSY–79	NLSY–97
Original NLSY sample size	3,874	3,455
Basic sample construction		
Keep age 14-17 in year 1979	55%	
Didn't drop in first 2 waves	96%	98%
Work history construction		
School-exit time identifiable	87%	99%
Out of school for 1+ years	90%	89%
Out of school for 8+ years	61%	68%
Adding explanatory variables		
Skill variables non-missing	99%	71%
Family variables non-missing	80%	87%
Neighborhood variables non-missing	97%	100%
School-to-work transition variables non-missing	98%	98%
Final sample size		
Black non-Hispanic men	271	396
White non-Hispanic men	444	825

[1] The original NLSY samples include Black and white non-Hispanic men (excluding the economically disadvantaged white subsample and the military subsample in NLSY–79).

[2] The percentage (%) in the table shows the share of the sample that remains after each line of sample choice, conditioning on all previous sample choices. For example, conditioning on previous sample choices, 99% of the remaining NLSY–79 sample has a complete list of skill variables and 71% of the remaining NLSY–97 sample has a complete list of skill variables.

Table A.2: Compare Final Sample with Original NLSY Sample

NLSY–79	Original Sample White men	Original Sample Black men	Final Sample White men	Final Sample Black men
Highest grade completed	13.6	12.6	13.4	12.9
AFQT percentile	55.7	22.1	57.0	22.9
Log parental income	9.2	8.9	11.5	10.8
Mother highest grade completed	12.1	11.0	12.1	11.1
Live with both parents	0.8	0.5	0.8	0.5
NLSY–97	**Original Sample White men**	**Original Sample Black men**	**Final Sample White men**	**Final Sample Black men**
Highest grade completed	13.9	12.5	12.8	11.8
AFQT percentile	54.3	27.5	49.2	25.2
Log parental income	11.1	8.6	10.9	8.6
Mother highest grade completed	13.4	12.4	12.9	12.3
Live with both parents	0.6	0.3	0.6	0.3

[1] The original NLSY samples include Black and white non-Hispanic men (excluding the economically disadvantaged white subsample and the military subsample in NLSY–79). The original NLSY samples are weighted by sample weights provided for wave one. The final samples, as in my main analysis, are weighted by NLS custom sample weights that account for the sample restrictions applied in my paper.

[2] The table presents separately for the original NLSY samples and my final samples of analysis the mean of individual and family variables. The upper panel is for the NLSY–79 and the lower panel is for the NLSY–97. See variable definitions in the main text.

Table A.3: DFL Sequential Decomposition (NLSY–97): Complete Results

NLSY–97	W-B Gap	Share Explained by			Residuals
Sequential Ordering I		Skill	Family	NBHD	
Weeks worked per year	6.90	27%	32%	-7%	48%
Log avg annual earnings	1.54	30%	18%	13%	39%
Sequential Ordering II		Family	Skill	NBHD	
Weeks worked per year	6.90	34%	25%	-7%	48%
Log avg annual earnings	1.54	20%	29%	13%	39%
Sequential Ordering III		Family	NBHD	Skill	
Weeks worked per year	6.90	34%	-6%	24%	48%
Log avg annual earnings	1.54	20%	9%	33%	39%

[1] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and NBHD stands for neighborhood. The sample is a balanced panel of 825 white men and 396 Black men who have completed formal schooling for at least eight years. Sample weights are used.

[2] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take the inverse hyperbolic sine.

Table A.4: DFL Sequential Decomposition (NLSY–97, Bootstrap Results)

NLSY–97	W-B Gap	Explained Gap			Residuals
Sequential Ordering I		Skill	Family	NBHD	
Weeks worked per year	6.90	1.86†	2.21†	-0.48	3.31
Log avg annual earnings	1.54	0.46†	0.28†	0.20§	0.60
Sequential Ordering II		Family	Skill	NBHD	
Weeks worked per year	6.90	2.35†	1.73†	-0.48	3.31
Log avg annual earnings	1.54	0.31†	0.45†	0.20§	0.60
Sequential Ordering III		Family	NBHD	Skill	
Weeks worked per year	6.90	2.35†	-0.41	1.67†	3.31
Log avg annual earnings	1.54	0.31†	0.14	0.51†	0.60

[1] This table present bootstrap results for 5,000 times. † (§) indicates a p-value below 0.05 (0.10). Note that the table shows the explained *gap* instead of the explained *share*.

[2] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and NBHD stands for neighborhood. The sample is a balanced panel of 825 white men and 396 Black men who have completed formal schooling for at least eight years. Sample weights are used.

[3] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take the inverse hyperbolic sine.

Table A.5: Is it Schooling or Measured Cognitive Skills? (NLSY–97): Complete Results

	W-B Gap	Share Explained by			Residuals
Skill includes only HGC		Skill	Family	NBHD	
Avg weeks worked per year	6.90	14%	22%	-7%	70%
Log avg annual earnings	1.54	12%	10%	9%	69%
		Family	Skill	NBHD	
Avg weeks worked per year	6.90	34%	3%	-7%	70%
Log avg annual earnings	1.54	20%	3%	9%	69%
		Family	NBHD	Skill	
Avg weeks worked per year	6.90	34%	-6%	1%	70%
Log avg annual earnings	1.54	20%	9%	3%	69%
Skill includes only AFQT		Skill	Family	NBHD	
Avg weeks worked per year	6.90	28%	31%	-8%	49%
Log avg annual earnings	1.54	30%	16%	13%	41%
		Family	Skill	NBHD	
Avg weeks worked per year	6.90	34%	25%	-8%	49%
Log avg annual earnings	1.54	20%	26%	13%	41%
		Family	NBHD	Skill	
Avg weeks worked per year	6.90	34%	-6%	23%	49%
Log avg annual earnings	1.54	20%	9%	30%	41%

[1] The top panel includes only highest grade completed (HGC) in the skill set, the bottom panel includes only Armed Forces Qualification Test (AFQT) score in the skill set. I use the AFQT score adjusted by Altonji, Bharadwaj, and Lange (2012).

[2] NBHD stands for neighborhood. The sample is a balanced panel of 825 white men and 396 Black men who have completed formal schooling for at least eight years. Sample weights are used.

[3] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take the inverse hyperbolic sine.

Table A.6: DFL Sequential Decomposition (NLSY–79): Complete Results

NLSY–79	W-B Gap	Share Explained by			Residuals
Sequential Ordering I		Skill	Family	NBHD	
Weeks worked per year	5.78	64%	18%	-8%	27%
Log avg annual earnings	1.37	67%	18%	-26%	41%
Sequential Ordering II		Family	Skill	NBHD	
Weeks worked per year	5.78	51%	30%	-8%	27%
Log avg annual earnings	1.37	67%	17%	-26%	41%
Sequential Ordering III		Family	NBHD	Skill	
Weeks worked per year	5.78	51%	59%	-37%	27%
Log avg annual earnings	1.37	67%	16%	-25%	41%

[1] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and NBHD stands for neighborhood. The sample is a balanced panel of 444 white men and 271 Black men who have completed formal schooling for at least eight years. Sample weights are used.

[2] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take the inverse hyperbolic sine.

Table A.7: DFL Sequential Decomposition (NLSY–79, Bootstrap Results)

NLSY–79	W-B Gap	Explained Gap			Residuals
Sequential Ordering I		Skill	Family	NBHD	
Weeks worked per year	5.78	3.70†	1.04	-0.46	1.56
Log avg annual earnings	1.37	0.92†	0.25	-0.36	0.56
Sequential Ordering II		Family	Skill	NBHD	
Weeks worked per year	5.78	2.95†	1.73	-0.46	1.56
Log avg annual earnings	1.37	0.92†	0.23	-0.36	0.56
Sequential Ordering III		Family	NBHD	Skill	
Weeks worked per year	5.78	2.95†	3.41†	-2.14	1.56
Log avg annual earnings	1.37	0.92†	0.22	-0.34	0.56

[1] This table present bootstrap results for 5,000 times. † (§) indicates a p-value below 0.05 (0.10). Note that the table shows the explained *gap* instead of the explained *share*.

[2] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and NBHD stands for neighborhood. The sample is a balanced panel of 444 white men and 271 Black men who have completed formal schooling for at least eight years. Sample weights are used.

[3] The number of weeks worked per year is averaged over the sixth to eighth years. The annual earnings are averaged over the sixth to eighth years and then take the inverse hyperbolic sine.

B Decomposition Method

In this Appendix, I describe the decomposition method (DiNardo, Fortin, and Lemieux (1996, hereafter DFL)) that I use for the main analysis of the paper. I first describe how the DFL method works under the context of explaining racial labor market gaps, and then discuss how to interpret the DFL estimates and relate them to other estimates in the literature.

B.1 Aggregate Decomposition

Let $f_w(y)$ be the density of labor market outcome y (such as employment or earnings) for white men and $f_b(y)$ for Black men. Let Z represent a vector of observed individual-, family-, and neighborhood-level characteristics that have an impact on one's labor market outcome y. The counterfactual density of y for white men who had the observed characteristics of Black men can be written as $f_w(y; Z_b)$. Intuitively, this counterfactual holds the *relationship* between y and Z as fixed for white men. The DFL method keeps this relationship non-parametric, so no specific functional form is imposed on $f_w()$.

Using this counterfactual, I can conduct the following decomposition of the racial gap in outcome y:

$$\begin{aligned} f_w(y) - f_b(y) &= f_w(\, y; Z_w \,) - f_w(\, y; Z_b \,) \\ &+ f_w(\, y; Z_b \,) - f_b(\, y; Z_b \,). \end{aligned} \tag{1}$$

The first line in Equation 1 represents the racial gap that can be explained by Black-white differences in observed characteristics Z (also known as "quantities"). The second line, which represents the unexplained residuals, include the contributions of 1) Black-white differences in *unobserved* characteristics and 2) Black-white differences in the *returns* (also known as "prices") paid to observed and unobserved characteristics. For example, if Black men receive lower labor market returns (such as less working time and/or lower wages) to the same characteristics due to discrimination, the racial differences in returns will be left in the unexplained residuals.[49]

[49]Equivalently, in principle one can conduct an alternative decomposition using $f_b(y; Z_w)$, the counterfactual outcome for Black men if they had the observed characteristics of white men. Conducting this reverse decomposition will introduce a common support problem, which has been discussed in earlier studies (Barsky et al., 2002; Heywood and Parent, 2012). Another distinction between the decomposition in Equation 1 and this reverse decomposition is whether $f_w()$, the earnings or employment function for white men, or $f_b()$, the function for Black men, is used. Under the context of racial labor market gaps, the literature usually uses $f_w()$ for decomposition analysis, mainly because the earnings or employment function received by white men is arguably more similar to

B.2 Sequential Detailed Decomposition

In addition to the *aggregate* decomposition, the DFL method allows me to estimate the contribution of different subsets of variables in Z to the racial gap in labor market outcome y. This *detailed* decomposition helps answer questions such as "What labor market outcomes would white men have achieved if they had the same family background and education and skills as Black men in the same cohort but kept their original childhood neighborhood characteristics and the school-to-work transition?"

Let Z consist of four main subsets of variables: family background F, childhood neighborhood N, education and skills S, and the school-to-work transition T. One of the possible detailed decompositions can be written as

$$\begin{aligned} f_w(y) - f_b(y) &= f_w(y) - f_w(y; F_b, N_w, S_w, T_w) \\ &+ f_w(y; F_b, N_w, S_w, T_w) - f_w(y; F_b, N_b, S_w, T_w) \\ &+ f_w(y; F_b, N_b, S_w, T_w) - f_w(y; F_b, N_b, S_b, T_w) \\ &+ f_w(y; F_b, N_b, S_b, T_w) - f_w(y; F_b, N_b, S_b, T_b) \\ &+ f_w(y; F_b, N_b, S_b, T_b) - f_b(y). \end{aligned} \tag{2}$$

The first line represents the contribution of Black-white differences in family background F. The contribution is the sum of a *direct* effect of family background F on labor market outcome y and an *indirect* effect, which comes from any changes in the distributions of N, S, and T that are attributed to the changes in F. In other words, this is the *unconditional* effect of family background on the racial gap in y.

The second line represents the contribution of Black-white differences in childhood neighborhood N *after* accounting for Black-white differences in family background characteristics. It is important to note that when holding family background constant between Black and white men, any variations in neighborhood characteristics that are implied by variations in family characteristics are also held to be constant between Black and white men. The third and fourth lines can be interpreted in a similar fashion as a conditional contribution of education and skills and the school-to-work transition, respectively. The last line represents the racial gap in y that

the hypothetical earnings or employment function in a labor market without discrimination (or other institutional barriers) against Black men. I therefore stick to the decomposition in Equation 1 throughout my analysis.

remains unexplained after accounting for Black-white differences in all observed factors in Z.[50]

An important feature of the DFL decomposition is that the detailed decomposition is not unique. As is shown in Equation 2, the contributions of different components of Z to the overall racial gap depend on the sequential ordering by which the different components (F, N, S, and T) are added into the decomposition. The components that are added earlier in the sequence are given more credit in explaining the racial gap. The merit of any sequential ordering depends on how the different components are causally related to the others. Under a similar context, Altonji, Bharadwaj, and Lange (2012) argue that a natural ordering is the one that follows the *timing* of variables.

In my empirical analysis, I explore different choices of sequential orderings as in a permutation exercise. Across all orderings, I always keep the school-to-work transition as the last component after all "pre-market" factors (F, N, and S), because one's transition performance is presumably an outcome of "pre-market" factors. Otherwise, I hold no prior as to where F, N, and S should be in the sequence relative to each other.

B.3 Interpreting DFL Decomposition Estimates

To help illustrate the DFL estimates and draw a more direct comparison to other estimates in the literature, in this section I first impose some additional structure on the DFL method following Altonji, Bharadwaj, and Lange (2012). Note that this section is for illustrative purpose only, and my main decomposition analysis does not require the additional structure discussed here.

Under linearity and additional separability assumptions, we can write down the relationship between outcome y and the underlying characteristics as

$$E(\,y; F, N, S, T\,) \quad = \quad \beta_0 + \beta_f F + \beta_n N + \beta_s S + \beta_t T, \tag{3}$$

where β is the partial effect of each underlying factor on y. This is the classical linear regression that many studies in the literature rely on.

To see how the DFL decomposition works, further assume that the relationship between the

[50] It is important to iterate that the DFL decomposition focuses on how much of the racial gaps in y can be explained by racial differences in N, F, S, and T ("quantities"), and it does not reveal the potential effect of racial differences in the returns paid to each one of these factors ("prices"), which will be absorbed in the residuals.

underlying factors is also linear and separable.[51] In a specific sequential decomposition, this assumption allows us to write the lower-order component as a linear and additive function of the higher-order components. For example, in the decomposition of Equation 2, we can write childhood neighborhood characteristics as a function of family background:

$$E(\,N;F\,) \;\; = \;\; \gamma_{0,n} + \gamma_{f,n}F, \tag{4}$$

where $\gamma_{f,n}$ is the partial effect of F on N. This equation only includes F on the right-hand side because F is the only factor added before N in the decomposition sequence.

Under these assumptions, the DFL sequential decomposition estimates can be written out *explicitly.* Take the decomposition of Equation 2 as an example. The contribution of family background F, which is added first in the sequence, can be expressed as

$$(\,F_w - F_b\,) \times (\,\beta_f + \gamma_{f,n}\beta_n + \gamma_{f,s}\beta_s + \gamma_{f,t}\beta_t\,). \tag{5}$$

The estimated contribution of F is the sum of two terms: 1) the partial effect of F on y, as represented by $(\,F_w - F_b\,) \times \beta_f$, and 2) an *indirect* effect on y that arises because shifts in F lead to shifts in N, S, and/or T, represented by $(\,F_w - F_b\,) \times (\,\gamma_{f,n}\beta_n + \gamma_{f,s}\beta_s + \gamma_{f,t}\beta_t\,)$.[52]

The contribution of childhood neighborhood N, which is added after F in the sequence shown in Equation 2, can be expressed as

$$(\,\widetilde{N}_w - \widetilde{N}_b\,) \times (\,\beta_n + \gamma_{n,s}\beta_s + \gamma_{n,t}\beta_t\,), \tag{6}$$

where $\widetilde{N}_w - \widetilde{N}_b = (\,N_w - N_b\,) - \gamma_{f,n}(\,F_w - F_b\,)$ is the *residualized* racial differences in N by F.

If racial differences in childhood neighborhood characteristics can be fully (or more than fully) accounted for by racial differences at the family level, the estimated contribution of N could be close to zero (or negative). In practice, it is not uncommon for DFL decomposition to produce estimates with a negative sign (e.g., DiNardo, Fortin, and Lemieux, 1996; Altonji, Bharadwaj, and Lange, 2012).

Similarly, the contribution of education and skills S, which is added after F and N in the

[51] See Altonji, Bharadwaj, and Lange (2012) for a detailed introduction of the assumptions.

[52] Altonji, Bharadwaj, and Lange (2012) call Equation 5 the "sequential marginal effect."

sequence, can be written as

$$(\widetilde{S}_w - \widetilde{S}_b) \times (\beta_s + \gamma_{s,t}\beta_t), \tag{7}$$

where $\widetilde{S}_w - \widetilde{S}_b = (S_w - S_b) - \gamma_{f,s}(F_w - F_b) - \gamma_{n,s}(N_w - N_b)$ is the *residualized* racial differences in S by F and N.

B.3.1 Relation to Other Estimates in the Literature

How does the DFL sequential decomposition estimate relate to other estimates in the literature? Here I discuss two types of estimates in particular, using the neighborhood effect literature as an example. The two estimates are 1) the estimated causal effect of neighborhoods and 2) the estimated contribution of neighborhoods to racial labor market gaps based on linear regressions.

First, a series of papers have estimated the causal effect of growing up in "good" neighborhoods (Chetty, Hendren, and Katz, 2016; Chyn, 2018; Chetty and Hendren, 2018a).[53] In Equation 6, the causal effect of neighborhood is represented by β_n. Conceptually, the contribution of neighborhoods estimated in the DFL decomposition therefore nests the causal estimates of neighborhood established in the literature, although β_n is not directly estimated in the DFL decomposition. An important implication from Equation 6 is that even when there is a causal neighborhood effect ($\beta_n > 0$), the explanatory power of neighborhood to the observed racial labor market gaps can still be limited if the *residualized* Black-white differences in childhood neighborhood characteristics are quantitatively small.

Second, an alternative approach to estimate the contribution of neighborhoods uses linear regressions (or decomposition methods based on linear regressions, e.g., Oaxaca-Blinder decomposition). One example of this approach is Chetty et al. (2020). Like the DFL decomposition, the estimates in this series of literature are also mainly descriptive. Using terminology in this section, the linear-regression-based estimates can be written as $(N_w - N_b) \times \beta_n$.

The DFL estimate in Equation 6 differs from the linear-regression-based estimate in two ways. The first and major distinction is that the DFL estimate uses the *residualized* racial differences in neighborhood characteristics $(\widetilde{N}_w - \widetilde{N}_b)$ rather than the raw differences $(N_w - N_b)$. As pointed out by Heckman (2018), to identify the true contribution of neighborhoods, it is

[53] The causal effect in this recent literature is usually estimated by comparing families (and children in these families) who move to "good" neighborhoods with families who stay in disadvantaged neighborhoods.

important to rule out the part of the raw racial differences in neighborhood characteristics ($N_w - N_b$) that reflect the residential sorting of Black and white families and individuals across neighborhoods. This is exactly what the residualized racial differences ($\widetilde{N}_w - \widetilde{N}_b$) help identify. The second distinction is that the DFL estimate includes the indirect effect that works through lower-order components in the sequence.[54]

B.4 Estimating the Counterfactual

The DFL method constructs the counterfactual $f_w(y; Z_b)$ by *reweighting* the joint distribution of (y, Z) for white men so that the *reweighted* distribution of Z for white men matches the distribution of Z for Black men. To see how the weight is determined, the counterfactual density $f_w(y; Z_b)$ is written as the following integral of the conditional density $f_w(y \mid z)$ over the Z distribution of Black men:

$$\begin{aligned} f_w(y; Z_b) &= \int f_w(y \mid z)\; dF_b(z) \\ &= \int f_w(y \mid z)\; \psi(z)\; dF_w(z), \end{aligned}$$

where the weight $\psi(z) = dF_b(z)/dF_w(z)$. Applying Bayes's rule, I rewrite the weight as

$$\psi(z) \;\; = \;\; \frac{dF_b(z)}{dF_w(z)} = \frac{Pr(z \mid b)}{Pr(z \mid w)} = \frac{Pr(b \mid z)}{Pr(w \mid z)} \frac{Pr(w)}{Pr(b)},$$

where $Pr(b \mid z)$ is the probability of being Black given on observed characteristics z and $Pr(b)$ is the unconditional probability of being Black. $Pr(b \mid z)$ can be estimated with a probit model that includes the full vector of z, and $Pr(b)$ can be estimated with the sample fraction of Black men. $Pr(w \mid z)$ and $Pr(w)$ can be estimated similarly. When estimating $Pr(b \mid z)$ and $Pr(w \mid z)$ with probit models, I impose parametric functional forms. Doing this makes the DFL method semi-parametric, not completely non-parametric.

Similar to propensity score matching, a practical issue in the DFL decomposition is how to deal with extremely large weights. Intuitively, the weight $\psi(z)$ will be large if the characteristics vector z is very rare among white men. In this case, $Pr(z \mid w)$ will be very small and $Pr(z \mid b)$

[54]Note that the DFL decomposition does not impose any functional form assumptions on the relationship between the outcome and the underlying factors, although in this section I impose linear and additive structures for illustrative purposes. This is another distinction between the DFL estimate and linear-regression-based estimates.

will be very large, which drives up the weight $\psi(z)$. In practice, I first adjust the weight $\psi(z)$ to have a mean of one and then cap the weight at the value of 20, under the prior that any weights above 20 should be due to sampling errors. What this capping does is basically *down-weight* white men who share similar observed characteristics z with Black men in the sample. By down-weighting these white men, the explanatory power of z to the racial gaps in y is also adjusted down.[55]

The counterfactuals in the detailed decomposition, as in Equation 2, can be estimated in a similar way. For example, write $f_w(\, y; F_b, N_b, S_w, T_w\,)$, the counterfactual density of y for white men when they had the same family background and childhood neighborhood as Black men, as the following integral:

$$\begin{aligned} f_w(\, y; F_b, N_b, S_w, T_w\,) &= \int f_w(\, y \mid f, n\, ;\, S_w, T_w\,)\; dF_b(f,n) \\ &= \int f_w(\, y \mid f, n\, ;\, S_w, T_w\,)\; \phi(f,n)\; dF_w(f,n). \end{aligned}$$

Using Bayes's rule, I can rewrite the weight $\phi(f, n)$ as

$$\psi(f,n) \;=\; \frac{dF_b(f,n)}{dF_w(f,n)} = \frac{Pr(b \mid f,n)}{Pr(w \mid f,n)}\frac{Pr(w)}{Pr(b)}.$$

As explained earlier, $Pr(b \mid f, n)$ and $Pr(w \mid f, n)$ can be estimated with a probit model that includes F and N as explanatory variables, and $Pr(w)$ and $Pr(b)$ can be estimated with the sample share of white and Black men. The same procedure can be applied to estimate other counterfactuals as well as the associated weights.

[55] Note that Altonji, Bharadwaj, and Lange (2012) also cap the weights in a similar context, where they reweight the NLSY–79 sample to make it similar to the NLSY–97 sample. My results are qualitatively robust to different choices of weight caps.