ABSTRACT

Title of dissertation: ESSAYS ON SKILLS AND RACIAL GAPS IN THE U.S. LABOR MARKET

Sai Luo, Doctor of Philosophy, 2020

Dissertation directed by: Professor Judith K. Hellerstein
University of Maryland Department of Economics

In this dissertation I establish some of the first evidence on the early career labor market experiences of young American men from the Millennial cohort. I also conduct a cross-cohort comparison of the early career outcomes of Millennials compared to their predecessors from the Baby Boomer cohort. The empirical analysis in this dissertation is facilitated by the 1997 and 1979 samples of the National Longitudinal Survey of Youth (NLSY).

First, I document the racial gaps in early career labor market trajectories of a cohort of early Millennial men (NLSY–97, born 1980–1984), and explore the driving forces behind them. Tracing the experiences of Black and white young men over their first eight years after school completion, I show that racial gaps in various labor market outcomes opened up immediately post-schooling, and largely persisted over the subsequent years. In particular, I find that measured Black-white disparities in accumulated education and skills, especially cognitive skills, play the central role in explaining the observed racial gaps in employment and earnings.

Second, I compare how the racial labor market gaps have changed between the Baby Boomers (NLSY–79, born 1957–1964) and these Millennials. Both Black and white men in the older cohort experienced upward-sloping trajectories in employment and earnings in the

first four to five years post-schooling. In the younger cohort, the labor market trajectories, especially for employment, were comparatively flatter both for Black men and for white men. Relative to the older cohort, a larger share of the racial employment and earnings gaps in the younger cohort cannot be explained by measured racial differences in observable pre-market characteristics. Yet education and skills remain the key explanatory factor among observable characteristics.

Third, in co-authored work, we examine how the wage returns to cognitive skills have evolved across cohorts of white men in the U.S. labor market. We show that the distribution of measured cognitive skills has diverged between the NLSY–79 and the NLSY–97. This divergence has a meaningful impact on estimated returns to cognitive skills. We explore why this divergence has occurred, considering both economic and measurement explanations, and we conclude that the conventional wisdom of a declining return to cognitive skills may well be incorrect.

ESSAYS ON SKILLS AND RACIAL GAPS IN THE U.S. LABOR MARKET

by

Sai Luo

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:
Professor Judith K. Hellerstein, Chair/Advisor
Professor Sergio Urzúa
Professor Jessica Goldberg
Professor Katharine Abraham
Professor Kevin Roy

Dedication

To Jingya. My girlfriend, my fiancée, and my wife.

Acknowledgments

I am deeply grateful to everyone who has helped me in my successful completion of this dissertation. It is impossible to thank each and every one of them.

First and foremost, I am indebted to my advisor Judy Hellerstein. She has held my hand and walked me through every stage of my graduate years, and I have become the researcher I am today because of her. I am grateful to Jessica Goldberg, her invaluable comments, and her constant encouragement. She once said Judy is a role model standing on her shoulder guiding her career. Now I have two standing on mine. I am grateful to Sergio Urzúa for countless discussions and critical comments that have pushed my research to another level. Aside from a close mentoring and advising relationship with all my advisors, I am proud of the work we have done together, and I am excited to continue working with them in my future career.

I would like to thank Katharine Abraham and Melissa Kearney for their insights and comments, which have helped polish my ideas clearer. I thank Kevin Roy for serving on my dissertation committee. During the writing of this dissertation, I have benefited enormously from discussions with Guido Kuersteiner, Seth Murray, Nolan Pope, Fernando Saltiel, Matt Staiger, Ben Zou, fellow graduate students at Maryland, and seminar participants at the Bureau of Labor Statistics. I gratefully acknowledge the financial support from the Economic Club of Washington D.C. and the Maryland Population Research Center.

On more personal notes, I would like to thank my mom and dad for all their moral support. I owe special thanks to Joon, for all the coffee, jogging, and Ping-Pong we have done together. The generous help from Vickie, Terry, Amanda, Mark, Heather, and other Maryland staffs has made my experience at Maryland smooth and easy.

Last, I would like to thank Jingya, who was my girlfriend when I started graduate program, and is my wife when I am about to finish. Without you, I would never have gotten where I am.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Millennials now make up more than one-third of the American labor force, a number projected to grow in years to come as older cohorts gradually leave the workforce (Pew Research Center, 2018; Bureau of Labor Statistics, 2019). In the 2018 midterm election, Millennials cast more than one-fifth of all votes, and their turnout nearly doubled from 2014 (Pew Research Center, 2019b).[1] In contrast to the increasingly notable role of Millennials in American economic, social, and political life, our understanding of the early career experiences of this cohort is still preliminary and relatively limited.

Growing anecdotal evidence suggests that Millennials are struggling to gain a foothold in the labor market and to climb up career ladders (The Atlantic, 2015; Forbes, 2016). This observation is worrisome not only because Millennials constitute a crucial part of the potential labor force but also because a failure to attach to the labor market may lengthen the time that it takes for Millennials to achieve other conventional milestones. Compared to earlier cohorts, such as the Baby Boomers, at similar ages, Millennials are more likely to live at the homes of their parents, are less likely to be married, and are less likely to have a child (Pew Research Center, 2017; 2020). Conducting a comprehensive and serious examination of the early career labor market patterns of the Millennials is then of great importance to public policies that focus on young Americans, both inside and outside the labor market.

[1] The Pew Research Center defines Millennials as the cohort born between 1981 and 1996. Baby Boomers are usually referred to as the cohort born between 1946 and 1964, based on the post-World War II birth surge.

This dissertation provides some of the first empirical evidence on the labor market experiences in the early careers of Millennial men and how the patterns have evolved across cohorts. In the second chapter, I examine the racial gaps in labor market outcomes among Black and white Millennial men and explore the underlying explanatory forces. The third chapter extends the analysis to compare the racial labor market gaps between the Millennial cohort and the Baby Boomer cohort. In the fourth chapter, coauthored with Judy Hellerstein and Sergio Urzúa, I study how the wage returns to cognitive skills have changed across cohorts in the U.S. labor market.

The empirical analysis in this dissertation relies mainly on the 1997 and 1979 cohorts of the National Longitudinal Survey of Youth (NLSY–97 and NLSY–79), two nationally representative samples of young Americans born in 1980–1984 and 1957–1964, respectively. The NLSY–97 cohort can be regarded as early Millennials, who were in their 30s in the 2015 survey, and the NLSY–79 cohort can be regarded as late Baby Boomers. The NLSY keeps a rich record of individual and family characteristics, and its restricted-use geocode files allow me to link to external sources to construct measures of neighborhood characteristics. This facilitates a comprehensive decomposition that consists of the main component of this dissertation's second and third chapter.

Since the Civil Rights Act of 1964, substantial economic progress has been made in closing racial gaps in various aspects of U.S. society (Smith and Welch, 1989). However, more recent evidence casts doubt on whether this relative Black progress is continuing (Hellerstein and Neumark, 2012; Neal and Rick, 2014; Council of Economic Advisors, 2016; Wilson and Rodgers, 2016; Daly, Hobijn, and Pedtke, 2017). For example, Bayer and Charles (2018) show that when the non-employed population is taken into account, the median earnings gap between Black and white men has grown since mid to late 1970s and has reached a level as large as it was in 1950.[2] Recent research shows that substantial and

[2]In particular, the incarceration rate has more than tripled since 1980, and criminal justice policies have

significant racial gaps still persist in the early career labor market outcomes among young Americans today (Chetty et al. 2020). In June 2020, demonstrations across the country in support of the Black Lives Matter movement has again brought to national attention the long-standing racial disparities in and beyond the labor market.

One key factor contributing to the persistent racial gaps in labor market outcomes is pervasive racial discrimination experienced by Black Americans in schools, during job searches, at workplaces, in interaction with the criminal justice system, in the housing market, and across numerous other facets of their lives (e.g. Massey and Denton, 1993; Lang and Lehmann, 2012; Reardon and Owens, 2014; Council of Economic Advisors, 2016). Past research suggests that having an arrest history generates future non-employment among young men and poor labor market prospects further induce criminal activity (Grogger, 1992; Grogger, 1998; Pager, 2003). Racial discrimination in the criminal justice system can therefore cause and enlarge the racial disparities in labor market outcomes. This discussion is especially relevant and particularly important in the context of the Black Lives Matter movement, because the Millennial cohort have spent much of their childhood and early adulthood in a period of historically high incarceration rates (Council of Economic Advisors, 2016).

In addition, education and skills accumulated prior to labor market entry have been shown in past studies to be a key determinant to labor market outcomes, health outcomes, and criminal behaviors (Heckman, Stixrud, and Urzúa, 2006; García, Heckman, and Ziff, 2019; García et al., 2020). Yet more than sixty years after *Brown v. Board of Education*, segregated schools are still preventing many Black children from obtaining education and skills that are crucial for their future success in work and life (Reardon and Owens, 2014;

been shifting toward more punitive treatment, the burden of which falls disproportionately more on Blacks (Neal and Rick, 2014; Council of Economic Advisors, 2016). Bayer and Charles (2018) show that ignoring the trend of increasing institutionalized population leads to an understatement of the racial earnings gap, especially since late 1970s.

The Atlantic, 2018; New York Times, 2019; Economic Policy Institute, 2020). Though the second and third chapter of this dissertation focus particularly on racial gaps in labor market outcomes, my empirical findings, as I will discuss later, are consistent with the academic literature and public discussion about the continuing Black disadvantages inside and outside the labor market.

Most of the existing academic narrative on racial gaps in the labor market comes from previous cohorts of Americans. The evidence to date on the drivers of the racial labor market gaps among Millennials is far from conclusive. 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; Castex and Dechter, 2014; Deming, 2017), one cannot simply assume that the early career experiences of previous cohorts apply to this new cohort of Americans.[3]

The second chapter documents the racial gaps in the early careers of Millennial men (NLSY–97 cohort) and evaluates the roles of different underlying factors behind the observed labor market gaps. I trace the trajectories of employment and earnings outcomes in the first eight years post-schooling and find that Black men fell substantially behind their white counterparts in the various labor market outcomes in the very first year post-schooling. Over the following years, the initial racial labor market gaps either stayed stable or grew even larger.

Using the semi-parametric decomposition method first introduced by DiNardo, Fortin, and Lemieux (1996), I estimate how much of the racial employment and earnings gaps observed over the sixth to eighth years post-schooling in the NLSY–97 cohort can be explained by measured racial differences in four factors: education and skills, family back-

[3] One of the first papers that compare the cohorts of Millennials and Baby Boomers is Altonji, Bharadwaj, and Lange (2012). When the authors wrote the paper, the sample of Millennials (NLSY–97 cohort) had not accumulated much labor market experience, and the focus of their paper is how the characteristics of young Americans have changed across cohorts.

ground, childhood neighborhood, and the school-to-work transition. I choose the sixth to eighth years because this is when employment and earnings outcomes of young men in the NLSY–97 cohort started to reach a relatively stable stage.

I first show that racial differences in education and skills play the key role in explaining the racial gaps in employment and earnings. The explanatory power of education and skills is mainly driven by measured racial gaps in cognitive skills rather than by gaps in formal schooling or non-cognitive and social skills. It is important to emphasize that cognitive skills can be accumulated through childhood. The measure of cognitive skills in the data, the Armed Forces Qualification Test (AFQT) score, is recorded around ages 12–18 and is at least partly a function of early childhood exposures to differential family investments and school influences (Cunha et al., 2006; Cunha and Heckman, 2007).[4] What my finding does suggest is that paying attention to the reasons for Black disadvantage in the skill accumulation process may be a promising pathway to reduce racial disparities in labor market outcomes. School segregation may have limited the opportunities for Black children to obtain education and skills, and to narrow the racial gap. As many studies have pointed out, Black men may have accumulated lower levels of education and skills as a result of anticipating discrimination in the labor market (Neal and Johnson, 1996; Heckman, 1998; Darity and Mason, 1998). If Black children and Black families anticipate that they will face discrimination in the labor market so that their skills will be rewarded unfairly, they

[4] An important question related to the AFQT score, just like other psychometric test scores for skills and abilities, is whether the test is biased. For the AFQT score, 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 dissertation is whether the score is racially biased. If so, the AFQT score will be partly capturing bias against Blacks in the test rather than true racial differences in cognitive skills. In 1991, the National Academy of Sciences (NAS) led a study in the military focusing on the racial fairness of the test 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 fairness of the test, as it links the AFQT score to *direct* measures of military job performance, which are rarely available in civilian datasets. Whether the findings of the NAS study can be applied to civilian population is largely an open question. In the literature, some studies cast doubt on the racial fairness of the AFQT score (Rodgers and Spriggs, 1996), while others come to the opposite conclusion (Heckman, 1998).

may underinvest in skills prior to labor market entry.

Focusing on a cohort close in age to the NLSY–97 cohort, Chetty et al. (2020) shows that racial differences childhood neighborhood, measured with census tract fixed effects, explain about 31% of the observed racial income gap in their data. When discussing this result, Heckman (2018) argues that it is unclear how much of the documented effect of neighborhood reflects the residential sorting of families and individuals across different locations rather than the effect of neighborhood per se. Using the NLSY–97 data, I show that unconditionally, my measures of childhood neighborhood characteristics explain 20%–30% of the racial earnings gap (and 10%–20% of the racial employment gap), which is arguably close to Chetty et al. (2020). When conditioning on racial differences in education and skills and family background, the explanatory power of measured childhood neighborhood characteristics falls to 7% for racial earnings gap (and goes away completely for the racial employment gap). This finding suggests that much of the documented unconditional role of neighborhood, both in the NLSY–97 cohort and in the cohort of Chetty et al. (2020), comes from one or both of two channels.

First, it is possible that the unconditional neighborhood effect on racial labor market gaps largely reflects the fact that people with different individual and family backgrounds live in different neighborhoods and that the true role of neighborhood is actually limited in the context of understanding racial gaps in labor market outcomes. Numerous articles have documented housing discrimination against Black Americans, even decades after the passage of the Fair Housing Act, and that housing discrimination has contributed to residential segregation by race and further harmed labor market prospects of Black workers (Kain, 1968; Massey and Denton, 1993; Yinger, 1995; Cutler and Glaeser, 1997; Ondrich, Ross, and Yinger, 2000; Charles, 2003; Zhao, Ondrich, and Yinger, 2006). The unconditional neighborhood effect that I document in the NLSY–97 data can therefore also be a result of racial discrimination in the housing market as reflected in neighborhood sorting.

Second, some of the racial gaps in education and skills may have originated from neighborhood influences. Though past studies have shown that growing up in good neighborhoods has a positive effect on children's educational outcomes (Aaronson, 1998; Chyn, 2018), it is unclear how large is the role of neighborhoods in terms of explaining the overall racial education and skill gaps. In the decomposition, I find that the explanatory power of education and skills is quantitatively robust no matter whether it is estimated conditioning on racial differences in childhood neighborhood characteristics or not. This suggests that the second channel is not well supported by the NLSY–97 data. In future research, I plan to further explore whether and how the racial gaps in education and skills have come from racial differences in neighborhood characteristics at a more local level (such as the census tract).

In addition to understanding Millennials alone, the third and fourth chapters compare them and their early career experiences with their predecessors in the Baby Boomer cohort. In the third chapter, I extend the analysis in the second chapter and ask how the early career racial gaps and the underlying driving forces have changed from the NLSY–79 cohort to the NLSY–97 cohort. Tracking the employment and earnings outcomes year by year for young men from the two cohorts, the starkest cross-cohort change is in the shapes of the employment and earnings trajectories. In the NLSY–79 cohort, both Black and white men experienced an upward-sloping trajectory in employment and earnings in the first four to five years post-schooling. The trajectory was steeper for Black men, and as a result, the racial gaps (especially in employment) narrowed over the early career years. In the NLSY–97 cohort, the labor market trajectories became much flatter for both races. In particular, from the first to the eighth year post-schooling, there was little to no improvement in employment outcomes at various margins.

Although the shapes of labor market trajectories have changed across cohorts, the change occurred for both Black and white men, and the racial gaps have not changed by a

statistically significant degree from the NLSY–79 cohort to the NLSY–97 cohort. To understand whether and how the underlying explanatory factors of the early career racial gaps have changed across cohorts, I again focus on racial gaps in employment and earnings outcomes observed over the sixth to eighth years post-schooling and apply the semi-parametric decomposition of DiNardo, Fortin, and Lemieux (1996).

I first look at the overall explanatory power of racial differences in pre-market characteristics, including education and skills, family background, and childhood neighborhood, measured with the same or similarly constructed variables from the two cohorts. In the NLSY–79 cohort, measured racial differences in all pre-market characteristics together account for about 70%–80% of the racial gaps in employment and earnings. In the NLSY–97 cohort, the explanatory power of pre-market characteristics falls to around 50%. These findings suggest that racial differences in unobservables are playing a bigger role in the younger cohort.

I discuss three cases of potential unobservables. First, I examine if the lower explanatory power of pre-market factors in the NLSY–97 cohort is due to omitting hard-to-measure, but important, neighborhood characteristics. Focusing solely on the NLSY–97 cohort, I incorporate extra variables that measure childhood neighborhood characteristics at a more detailed level. Although these variables alone have meaningful explanatory power, including them adds little extra explanatory power to the existing set of pre-market characteristics. Second, I look at the role of the school-to-work transition, which has been documented to have a persistent impact on later outcomes (e.g. Kahn, 2010; Schwandt and Wachter, 2019). Conditioning on racial differences in pre-market characteristics, I show that Black disadvantage in the school-to-work transition, measured as weeks worked in the first year post-schooling or county unemployment rate in the year of labor market entry, can explain an additional 10% of the racial gaps in employment and earnings at the sixth to eighth years post-schooling. Third, I discuss the channels through which labor market discrimination

against Black men could affect my results. For example, it is well documented in past studies that Black men face discrimination in the hiring process of entry level jobs (summarized in Lang and Lehmann, 2012). During the school-to-work transition, Black men may end up with worse employment outcomes because of discrimination, or they may be discouraged from actively searching for jobs as a response to the expectation of discrimination. I argue that understanding how racial discrimination has evolved overtime in the U.S. labor market merits more future research.

From the NLSY–79 cohort to the NLSY–97 cohort, education and skills have remained as the key explanatory factor among all observable pre-market characteristics, confirming the findings of Neal and Johnson (1996) and the second chapter. But the absolute explanatory power of education and skills has fallen across cohorts, as the racial gap in cognitive skills, measured by the AFQT score, has narrowed substantially across cohorts. In the NLSY–79 cohort, the explanatory power of education and skills almost fully comes from measured racial differences in family background and childhood neighborhood characteristics. In contrast, in the NLSY–97 cohort, the role of education and skills in explaining racial labor market gaps is largely independent of racial differences in family and neighborhood characteristics, as measured in the NLSY data.

The third chapter examines how pre-market characteristics have changed across cohorts and how they have affected racial labor market gaps. In addition to changing characteristics, changing labor market returns to the same characteristics could have also led to different early careers outcomes of young men from the two cohorts. In the fourth chapter, I study how the labor market return to one of the key pre-market characteristics, cognitive skills, has changed from the NLSY–79 cohort to the NLSY–97 cohort.

The structure of the U.S. labor market has changed dramatically over the past few decades, potentially favoring some types of skills over others (Autor and Dorn, 2013; Beaudry, Green, and Sand, 2016; Autor, 2019). It is challenging to establish how labor

market returns to a specific type of skills at the individual level have evolved over time, as doing so requires a measure of skills that is consistently observed over time. Two influential studies provide empirical evidence that the wage returns to cognitive skills have declined from the NLSY–79 cohort to the NLSY–97 cohort, while the wage returns to other types of skills have increased (Castex and Dechter, 2014; Deming, 2017). Specifically, both studies estimate Mincerian wage regressions for the two NLSY cohorts and find that the OLS estimates of returns to cognitive skills, measured by the AFQT score, have gone down across cohorts. Using the empirical evidence they make from the NLSY data, both studies try to further explore why the returns have changed and conclude that technological change has fundamentally shifted the relationship between skills and labor market outcomes over the past 40 years.

In the fourth chapter, my coauthors and I re-examine the finding of declining returns to cognitive skills among the sample of white men. We first show in a simple investment model that if technology is indeed driving the changing returns to cognitive skills, it should also be driving investments in cognitive skills and therefore the population distribution of cognitive skills, a theoretical implication that the previous two studies have not formally taken into account. Applying the non-parametric method introduced in Yitzhaki (1996), we show that the Ordinary Least Squares estimates of how the returns to cognitive skills have changed across cohorts, as established in previous studies, are critically dependent on the changes in the distribution of the measured cognitive skills (AFQT score). In particular, there is a strong divergence in the distribution of the AFQT score from the NLSY–79 cohort to the NLSY–97 cohort: while a greater share of young Americans in the NLSY–97 cohort achieve high AFQT scores than in the NLSY–79 cohort, a greater share of the younger cohort are also being left behind with very low AFQT scores.

Non-parametrically, we decompose the Ordinary Least Squares estimate of wage returns to cognitive skills into a weighted average of unit treatment effects. The weighting

function is purely governed by the distribution of the AFQT score. As the distribution of the AFQT score diverged from the NLSY–79 cohort to the NLSY–97 cohort, the Ordinary Least Squares estimates are weighting the unit treatment effects differently in one cohort than in another. To make the estimates directly comparable between the two cohorts, we hold the weighting function consistently fixed at the NLSY–79 level for both cohorts and show find that the estimated returns to cognitive skills actually slightly increase from the NLSY–79 cohort to the NLSY–97 cohort.

Have the wage returns to cognitive skills really declined in the U.S. labor market? We argue that the jury is still out. Evidence suggests that the AFQT score in the NLSY–97 cohort may be subject to non-classical measurement errors, which adds further complications to a cross-cohort comparison (Schofield, 2014). However, this does not exclude the possibility that there has been a real decline in returns to cognitive skills. In fact, we find some suggestive evidence that if the decline is real, it is likely not happening uniformly across the whole cognitive skill distribution. We construct local estimates of the relationship between log wages and AFQT scores and find that the relationship is relatively flat in the NLSY–97 cohort for people with below-median AFQT scores. In particular, the "decline" in the estimated wage returns to cognitive skills from the NLSY–79 cohort to the NLSY–97 cohort seems to be solely driven by people with below-median AFQT scores. There is not a similar "flatness" among the sample of white women in the NLSY–97 cohort, suggesting measurement error may not be the sole driver of the story and that the declining returns to cognitive skills may be real for low AFQT scorers.

The focus of this chapter is on cognitive skills, but what about other types of skills, such as non-cognitive skills or social skills? Compared to the measurement of cognitive skills, there is less consensus among economists (and social scientists in general) on how to measure non-cognitive or social skills. Deming (2017) draws information from different sections of the NLSY–79 data and the NLSY–97 data to measure non-cognitive and social

skills and shows that the estimated wage returns to non-cognitive skills, and especially social skills, have gone up across cohorts. Given that the Ordinary Least Squares estimate of returns to skills is highly dependent on changes in the skill distribution, there is legitimate concern about the measurement of non-cognitive and social skills and how the distributions of non-cognitive and social skills have evolved across cohorts. Future research needs to also re-examine the influential finding of Deming (2017) and construct alternative non-cognitive and social skill measures that are more comparable across cohorts (and across datasets).

This dissertation aims to paint a broad picture of the early career labor market experiences of Millennial men and how they have evolved relative to a previous cohort of Baby Boomers. The second and third chapters focus on racial gaps in labor market outcomes of young men and show that, in both cohorts, education and skills play the key role in explaining the observed racial labor market gaps. The third and fourth chapters make cross-cohort comparisons, with different emphases on the pre-market characteristics of young men and labor market returns to the characteristics (with a special focus on cognitive skills). The findings in this dissertation add new evidence to our understanding of the importance of skills in the U.S. labor market.

Chapter 2: Understanding Racial Gaps in the Early Careers of Millennial Men

2.1 Introduction

More than five decades after the Civil Right Act was signed into law, racial gaps persist in various dimensions of the U.S. labor market. Although substantial economic progress was made in closing these racial gaps (James Smith and Welch, 1989), the trends seem to have stagnated or even reversed since 1980 (Wilson and Rodgers, 2016; Daly, Hobijn, and Pedtke, 2017). As the cohort of Millennials comprise an increasingly important share of the American labor force, recent evidence has also documented substantial income gaps between Black and white men in this cohort (Chetty, Hendren, Jones, and Porter, hereafter CHJP, 2020).[1]

Most of the existing narrative on the causes of racial gaps in the labor market comes from previous cohorts of Americans. The evidence to date on the drivers of the racial labor market gaps among Millennials is far from conclusive. Do Black-white differences in skills play an important role, as with previous cohorts? What about the roles of family background, childhood neighborhood, the school-to-work transition, and discrimination?[2]

[1]The Pew Research Center defines Millennials as the cohort born between 1981 and 1996 (Pew Research Center, 2019a). Millennials now make up more than one-third of the American labor force, a number projected to grow in years to come as older cohorts gradually leave the workforce (Pew Research Center, 2018; Bureau of Labor Statistics, 2019).

[2]The literature on the potential determinants of racial labor market gaps is enormous, and many of the papers are based on previous cohorts of Americans. For example, pre-market skills (or human capital) have been shown to be crucial in understanding racial gaps in labor market outcomes (Neal and Johnson, 1996;

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; Castex and Dechter, 2014; Deming, 2017), one cannot simply assume that the early career experiences of previous cohorts apply to this new cohort of Americans.

This chapter answers these questions by evaluating the roles of different factors in shaping the racial gaps in the early career experiences of Millennial men.[3] I study the 1997 cohort of the National Longitudinal Survey of Youth (NLSY–97), a nationally representative sample of young Americans born between 1980 and 1984, and document the racial gaps among young men in this cohort in their labor market trajectories, observed in the first eight years beyond schooling completion. In both employment and earnings, Black men lagged substantially behind their white counterparts beginning in the first year out of school. The racial gaps in employment and earnings largely persist in the following early career years. This persistence motivates my decomposition analysis, in which I explore what has driven the observed racial gaps in this cohort.[4]

In particular, I study the contributions of individual skill, family background, childhood neighborhood, and the school-to-work transition. I harness the richness of the NLSY–97 and its restricted geocode file to include a detailed list of observable characteristics in each

Heckman, Stixrud, and Urzúa, 2006; Urzúa, 2008). Family background and parenting style have been long understood as pivotal predictors for children's outcomes, and evidence shows that there are important racial differences in how parents raise and educate their children (Lareau, 1987; McAdoo, 2002; Thompson, 2018). "Good" childhood neighborhoods are shown to have an impact both on future adult outcomes (Aaronson, 1998; Chetty, Hendren, and Katz, 2016; Chyn, 2018; Chetty and Hendren, 2018a) and on reducing racial gaps in adult outcomes (CHJP, 2020), although it remains largely unknown what constitutes a "good" neighborhood. School-to-work transitions (i.e. how people initiate their careers) are found to have a persistent impact on future labor market outcomes (Light and Ureta, 1995; Neumark, 2002; Kahn, 2010; Rothstein, 2019; Rinz, 2019; Schwandt and Wachter, 2019; Yagan, 2019). The racial difference in school-to-work transition has been less explored. In addition, an important series of studies has emphasized the role of discrimination (e.g. Donohue and Heckman, 1991; Pager, 2003; Bertrand and Mullainathan, 2004; Charles and Guryan, 2008; Council of Economic Advisors, 2016).

[3]I focus on the racial gaps between Black and white non-Hispanic men. There are other important racial and ethnic gaps among both men and women that merit exploration in future research.

[4]Throughout this chapter, I use the word "cohort" to refer to men born between 1980 and 1984, not a specific birth year.

of these four factors.[5] Applying the semi-parametric decomposition method introduced by DiNardo, Fortin, and Lemieux (1996), I establish two key findings regarding the extent to which the four factors have driven the racial employment and earnings gaps observed in this cohort of young men.

First, racial differences in measured individual skill explain up to half of the mean racial gaps in employment and earnings. This central role of skill is attributable primarily to racial differences in measured cognitive skills rather than to gaps in formal schooling. Looking at racial gaps at the 25th percentile, the median, and the 75th percentile of Black and white men's employment and earnings distributions, individual skill differences also usually have the largest explanatory power.

Second, on its own, the set of childhood neighborhood characteristics I observe explains approximately 10%–20% of the mean racial employment gap and approximately 20%–30% of the racial earnings gap. Conditional on family background and individual skill, however, the explanatory power of neighborhood characteristics is negligible. Although the geocode file of NLSY–97 does not allow me to control for neighborhood characteristics that are as detailed as in some other datasets (such as in tax records), my neighborhood measures unconditionally achieve explanatory power close to what census tract fixed effects do under a similar context in a different study (CHJP, 2020).[6]

The central role of skills in explaining racial labor market gaps is consistent with numerous existing studies conducted on previous cohorts (Neal and Johnson, 1996; Heckman, Stixrud, and Urzúa, 2006; Urzúa, 2008). For example, in Neal and Johnson (1996)'s influential work, cognitive skills, as measured by the Armed Forces Qualification Test (AFQT)

[5] I largely follow the literature in variable construction for individual skill and family background (Altonji, Bharadwaj, and Lange, 2012; Deming, 2017; Baumrind, 1991; Doepke and Zilibotti, 2017). I use the restricted geocode data of the NLSY–97 and control for childhood neighborhood characteristics at as detailed a level as the data allow, and I measure the school-to-work transition with weeks worked during the first year beyond schooling completion.

[6] The geocode file of NLSY–97 is as detailed as the county level, and I also include some measures of within-county neighborhood quality.

score, on their own account for about 60% of the racial wage gap among young men from the NLSY–79, an older cohort born between 1957 and 1964.[7] Urzúa (2008) makes a distinction between measured cognitive skills (the AFQT score) and underlying cognitive ability, and shows that cognitive ability explains about 40% of the racial gaps in wages and earnings in the NLSY–79.[8]

This chapter provides the first evidence on the key importance of skills in understanding racial labor market gaps for the cohort of Millennials. Until now, our understanding of the sources of racial labor market gaps in this cohort largely has been shaped by the recent paper of CHJP (2020), who emphasize the key role of childhood neighborhood in explaining the racial income gap observed in their data. My findings do not directly contradict their findings, as CHJP do not incorporate direct measures of cognitive skills into their analysis. However, my finding that the explanatory power of childhood neighborhood characteristics diminishes to small or non-existent suggests that either one of the following two stories, or both. First, the unconditional explanatory power of neighborhood documented both in the NLSY–97 and in CHJP's data may reflect residential sorting of Black and white families into different neighborhoods, as pointed out by Heckman (2018). Second, if there is indeed

[7] An important question related to the AFQT score, just like almost all other psychometric test scores for skills and abilities, is whether the test is biased in favor of one group over another. For the AFQT score, since its first introduction by the Department of Defense for screening enlistees and assigning them to different occupations, a key question especially relevant for the purpose of this chapter 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 racial fairness of the test 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 fairness of the test, as it *directly* observes and measures military job performance and links it to the AFQT score, which is hardly available in civilian datasets. Whether the findings of the NAS study can be applied to civilian population is largely an open question. In the literature, some studies cast doubt on the racial fairness of the AFQT score (Rodgers and Spriggs, 1996), while others conclude otherwise (Heckman, 1998).

[8] In a structural model, Urzúa (2008) emphasizes the key insight that observed (AFQT) test score is a function of both underlying (cognitive) ability and other characteristics, including family background characteristics (such as parental income). According to the model, cognitive ability explains a smaller share of the racial gaps in labor market outcomes than AFQT score, because the racial gap in the AFQT score also picks up racial differences in family background characteristics that have a direct effect on labor market outcomes. Although I do not formally model the skill formation process, my interpretation of the estimated explanatory power of the AFQT score is consistent with the intuition of Urzúa (2008).

a true effect of neighborhoods on racial labor market gaps, it is likely working through the channel of skill formation.

What does the primary role of racial skill differences imply? Although there is no policy panacea to reduce racial labor market gaps, my findings shed light on potentially promising pathways as we move forward. In particular, my findings reinforce older studies that emphasize the critical role of skill development and suggest that it is important to continue to focus on institutional and economic barriers to Black men in the skill accumulation process.

For example, the measure of cognitive skills in my data, the AFQT score, is observed at ages 12–18, and is a function of a series of family investments and neighborhood influences in early childhood years. It is possible that some of the racial differences in measured cognitive skills in my data could have originally come from Black and white men's earlier childhood exposure to different *unobserved* neighborhood characteristics, such as local school quality. Meanwhile, as emphasized in Cunha et al. (2006), family investments may play a more crucial role than schools in children's skill accumulation process. Identifying the specific mechanisms behind the racial skill differences among Millennials is essential to understanding the racial gaps in the labor market outcomes of this and future cohorts.

An important source of the observed racial gaps in the U.S. labor market is discrimination (e.g. Donohue and Heckman, 1991; Pager, 2003; Bertrand and Mullainathan, 2004; Charles and Guryan, 2008; Council of Economic Advisors, 2016), and there are multiple channels through which discrimination relates to the factors incorporated in my analysis. First, as has been widely discussed, discrimination can have a "feedback" effect on individuals' pre-market investment decisions. For example, Black men (or their parents) who anticipate that there will be labor market discrimination may underinvest in skill accumulation.[9] Second, racial differences in the school-to-work transition also could be picking

[9]Discrimination beyond the labor market, such as in the criminal justice system, can have a similar feed-

up discrimination faced by Black men in their initial labor market experiences. Third, 21% of the mean racial employment gap and 7% of the mean racial earnings gap remain unexplained in my data. If, due to discrimination, Black and white men are receiving different labor market returns to the same characteristics, then this will be reflected in the residuals. Note that the impact of discrimination on racial labor market gaps goes beyond the border of 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 Black 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.

The remainder of this chapter proceeds as follows. In Section 2, I describe the NLSY–97 data and present descriptive facts on the racial gaps in early career work trajectories of young men in this cohort. Section 3 describes the semi-parametric decomposition method. Section 4 details the definition of individual skill, family background, childhood neighborhood, and the school-to-work transition used in the decomposition and summarizes racial differences in these characteristics. Section 5 presents the decomposition results. In Section 6, I summarize the main findings and lessons.

2.2 Description of Racial Gaps in Early Careers of the NLSY–97

2.2.1 NLSY–97 Data: A Sample of Early Millennials

The primary dataset used in this chapter is the NLSY–97, a nationally representative sample of Americans born between 1980 and 1984. According to the Pew Research Cen-

back effect. For example, if black men anticipate that racial discrimination in the criminal justice system will increase the probability that they will have a criminal record that, in turn, will harm their future labor market prospects (e.g. Pager, 2003), they may invest less in skills that will be rewarded in the labor market.

ter's definition, this NLSY cohort can be considered early Millennials. The respondents were between age 12 to 16 at the first interview in 1997, and they continue to be interviewed on an annual or biannual basis. My sample of analysis includes Black and white men from the core sample and the supplement minority sample of the NLSY–97. I use the NLSY custom sample weights both when creating summary statistics and when conducting decomposition analysis.

As the whole NLSY–97 cohort is now in their 30s (as of the most recent survey in 2015), the vast majority should have completed their formal schooling and should have had the opportunity to participate in the labor force for a number of years. Thus, now is the appropriate time to study the early career experiences of this cohort of young Americans without serious concern of sample truncation. Defining the exact schooling completion time is challenging because most datasets do not keep a detailed record of individuals' school enrollment history and young adults often move back and forth between school and work in their early careers. The NLSY–97 has a monthly retrospective diary of school (including college) enrollment status, which allows me to identify the exact time that an individual stops enrolling.

I follow the literature in defining schooling completion and work trajectories (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 completion. In my preferred sample, I keep a balanced panel of young men who completed schooling at least eight years (or 96 months) prior and track their labor market outcomes through the first eight years post-schooling. As I show in the next section, the work trajectories of both Black and white men in the NLSY–97 reach a relatively steady stage about six to eight years beyond school completion. I also show robustness using an alternative unbalanced panel, which includes up to eight years of labor market experiences for young men who completed school at least two years prior. My preferred balanced panel

includes 839 white men and 406 Black men, and the alternative unbalanced panel includes 1,210 white men and 534 Black men. [10]

For early Millennials in the NLSY–97, an important and distinctive marker of their early adulthood is the Great Recession. Rinz (2019) shows that compared to previous cohorts, the Millennials suffered more from the Great Recession in terms of a greater employment loss and a more long-lasting earnings loss.[11] Regarding my focus on racial gaps in early career experiences, if Black and white men left school and entered the labor market at different times, the Great Recession may have different impacts on their labor market outcomes. I examine this cohort's exposure to the Great Recession in Figure 2.1, which separately plots the corresponding calendar years for the first and eighth year post-schooling for Black and white men in the NLSY–97. In my sample, the vast majority of this cohort had already completed schooling before the start of the Great Recession in late 2007, and the first post-schooling year is spread generally evenly from 1997 to 2008.[12] Some in this cohort who left school earlier experienced the entire eight post-schooling years prior to the Great Recession, and those who left later spent at least some of the eight years under the shadow of the Great Recession.

[10]My definition of schooling completion and work trajectories involves two steps. First, I define young men as having completed schooling in a given month-year if they are not enrolled in school in any month of the year and in any following years in the sample. For example, if a young man graduated from high school, worked for a few years, went back to college, and rejoined the workforce later, their post-schooling experiences are defined to only include the post-college years. This definition therefore excludes two kinds of work experiences: 1) part-time jobs while enrolled in school and 2) relatively temporary work spells that are followed by returning back to school (as in the previous example). These short-term work experiences are not the focus of my analysis here but might be of particular interest to other research purposes. Second, I restrict the sample to a *balanced* panel of young men who have completed formal schooling for at least *eight* years and track their labor market outcomes through the first eight years post-schooling. The two costs of requiring a balanced panel of eight years are sample size (as some young adults have only completed schooling for less than eight years) and potential sample selection (as some young adults may change their school-completion time based on their anticipation of labor market prospects). My findings are robust to using an alternative *unbalanced* panel, which includes up to eight years of work experiences for young adults who have completed schooling for at least *two* years.

[11]Rinz (2019) defines the Millennials as born between 1981 and 1996 and compares them to three previous cohorts (Generation X, Baby Boomers, and the Silent Generation).

[12]In the alternative unbalanced panel, the first post-schooling year spreads from 1997 to 2014, and the Black-white difference in the timing of school completion is also modest.

Arguably, the more important pattern in Figure 2.1 is that the Black-white difference in the timing of school completion is modest. On average, Black men left school and entered the labor market earlier than white men, which is not surprising given that Black men had lower education levels on average, but the difference is limited and does not seem to be concentrated in any specific calendar year.[13] If anything, this modest difference indicates that Black men's early careers are slightly *less* exposed to the Great Recession, on average.[14] Even though Black and white men in the NLSY–97 entered the labor market at similar times, it is still possible that they experienced differential exposures to the Great Recession because they lived in different geographic areas or worked in different occupations and industries. As I introduce in detail in Section 3, I incorporate a measure of the school-to-work transition, which is one's employment status in the first year post-schooling, in my decomposition analysis. To the extent that the differential impact of the Great Recession is reflected in Black and white men's employment immediately beyond schooling, my school-to-work transition measure absorbs the potential effect of the Great Recession.[15]

2.2.2 Descriptive Facts on Racial Gaps in Early Career Trajectories

How do Black and white men of this cohort fare in their early careers? What are the racial gaps in their very first year beyond school completion, and how do the initial racial gaps evolve over time? In this section, I summarize the labor market trajectories of Black and white men in the NLSY–97. I start by visually presenting how employment and earnings evolve from the first through eighth year post-schooling.

[13] On average, Black men were also slightly younger than white men in the first post-schooling year, but the difference is small in magnitude (19.3 years old for Black men and 19.6 years old for white men).

[14] One thing to note is that for Black men, a visibly larger share of the eighth year corresponds to the start of the Great Recession around 2007–2009. This might explain why some employment and earnings trajectories for Black men, as presented in Figures 2.2–2.4, show a small dip at the eighth year. However, my decomposition results are robust to excluding the eighth year from the analysis.

[15] Otherwise, the effect of the Great Recession on the racial labor market gaps will be captured by the residuals in the decomposition analysis.

Figures 2.2–2.3 plot the trajectories of employment outcomes at different margins, and Figure 2.4 plots the trajectories of annual earnings.[16] The consistent pattern observed across outcomes is the substantial racial gap that began immediately in the first year post-schooling and the gap's strong persistence throughout the eight early career years. For example, 96% of white men were employed in their first year beyond school completion, while fewer than 85% of Black men were employed. This 11 percentage point gap narrows somewhat over the following years but largely persists and is statistically significant throughout the eight years.

At the intensive margin of employment, white men on average, worked for 41 weeks in their very first year post-schooling, while Black men, on average, worked for 32 weeks. Again, this large and significant gap persists to the eighth year. Similar differential trajectories are documented for employment outcomes at other margins, such as the share of people who worked for at least 26 weeks a year and the share who worked for at least 50 weeks a year (which can be thought of as a measure of "half-year" and "full-year" employment, respectively). As shown in Figure 2.4, as with employment, substantial initial racial gaps are also observed in both annual earnings and log annual earnings in the first year beyond schooling completion. The initial gaps either stay largely stable (for log earnings) or grow (for earnings) through the early career years.

The strong persistence of the racial employment and earnings gaps suggests that to understand the gaps, it is essential to pay particular attention to pre-market factors (such as skills that young men are about to bring to the labor market) *and* the events affecting Black and white men immediately upon schooling completion and labor market entry.[17] This pattern motivates my focus on individual skill, family background, and childhood

[16] Annual earnings are adjusted to 2013 dollars. I use inverse hyperbolic sine to allow for zeroes. For simplicity, I use the word "log" instead of inverse hyperbolic sine throughout the chapter

[17] It is possible that the work trajectories of Black and white men in the NLSY–97 will either converge or further diverge in the longer run. This goes beyond the scope of this chapter, but it is something I could investigate when the NLSY–97 cohort are further into their 30s and 40s.

neighborhood characteristics (which are presumably pre-market factors) and the school-to-work transition (measured in the first year post-schooling) when exploring what has driven the documented racial gaps in the decomposition analysis. Since the school-to-work transition is partly an outcome, it helps to capture the effect of unobserved pre-market factors and the effect of unobservables in the labor market, such as discrimination against Black men and differential exposure to the Great Recession between Black and white men, *as long as* the unobservables are reflected in young men's initial labor market experiences.

Table 2.1 summarizes Black-white gaps in their early career experience at the initial stage (the first year), at the relatively stable stage (the sixth to eighth year), and over the entire early career of eight years. On average, white men began their first post-schooling job at 11 weeks, while it took Black men 34 weeks to be employed. This 23-week gap means that compared to their white counterparts, an average Black man went through an immediate non-employment spell right after school completion that was more than five months longer.[18]

This difficulty faced by Black men in getting a foothold in the labor market is extremely worrisome. Historically, most job mobility and wage growth happens in the first few years of one's career (Topel and Ward, 1992). If Black men are disconnected from employment during this period, their career progression is delayed and possibly harmed permanently (Kahn, 2010; Kondo, 2015; Schwandt and Wachter, 2019). Additionally, past research suggests that poor labor market prospects induce criminal activity among young men and having an arrest history generates future non-employment (Grogger, 1992; Grogger, 1998). Considering the historically high incarceration rate over childhood and early adulthood of this cohort, failing the school-to-work transition may have a particularly large and long-

[18]This measure of the school-to-work transition duration is constructed over the first eight years post-schooling. For young adults who were never employed throughout all eight years, their transition duration is capped at eight years and is therefore underestimated. In my sample, 2% of Black men and less than 1% of white men were never employed in the entire eight years, suggesting the actual racial gap in transition duration is likely larger than 23 weeks.

lasting impact on Black men who did not immediately find a job and subsequently became involved in criminal activity.[19]

Because Black men are less attached to employment, it is well documented that simply comparing the earnings of *employed* Black and white men will likely underestimate the true racial earnings gap (Heckman, Lyons, and Todd, 2000; Johnson, Kitamura, and Neal, 2000; Neal, 2004; Bayer and Charles, 2018). Table 2.1 confirms this pattern by looking at annual earnings versus annual earnings excluding zeroes. In the first year post-schooling, when zero earnings are excluded, average earnings of white men are 46% higher than Black men (a log earnings gap of 0.46). When zero earnings are included (as in Figure 2.4), the racial gap in the first year increases to 215% (a log earnings gap of 2.15). When I take the average of annual earnings over the sixth to eighth years, white men earn 52% higher earnings when zeroes are excluded and 145% when zeroes are included. To incorporate this important Black-white difference in attachment, I always include zero earnings in the decomposition analyses.[20]

Given the unstable nature of the first few years of one's career, for the decomposition analysis, I primarily focus on racial gaps in labor market outcomes measured at the relatively later stage. Specifically, I take the average weeks worked per year and the average annual earnings over the sixth to eighth years post-schooling. The racial gaps in these outcomes are summarized in the second panel of Table 2.1. As the work trajectories in Figures 2.2–2.4 show, both employment and earnings see more growth and fluctuations in the first few years and mostly stabilize around the sixth to eighth year. Taking the average

[19]In results not presented here, I find that in the first eight years post-schooling, 41% of Black men in the NLSY–97 have at least one arrest and 17% of Black men have at least one episode of incarceration. The corresponding numbers are 25% and 7% for white men in the NLSY–97. The racial gaps also exist along the intensive margin (e.g., number of arrests, months of incarceration) and are all statistically significant. My decomposition analysis to date has been focusing on racial gaps in labor market outcomes. As I proceed, I may incorporate criminal involvement as an additional outcome.

[20]In results not presented here, I impute missing earnings values either with zeroes or based on broad earnings categories asked in the survey, and show that the basic patterns of racial gap hold.

over three years, instead of looking at a single year (such as the eighth year), also reduces potential measurement errors. As a robustness check, in the next section I also present decomposition results focusing on racial gaps averaged over a longer period, from the second to eighth year.

In sum, the early careers of young men in the NLSY–97 are characterized by a substantial and persistent racial gap in various employment and earnings measures. In the following section, I apply semi-parametric decomposition methods to explore the drivers of the documented racial labor market gaps in the NLSY–97.

2.3 Decomposition Method

I now describe the method that I apply to assess the contribution of individual skill, family background, childhood neighborhood, and the school-to-work transition to the documented racial employment and earnings gaps in the NLSY–97. 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; CHJP, 2020).[21] However, there is evidence showing that some of the parametric assumptions widely imposed in classical regression specifications are not supported by the data (Heckman, Stixrud, and Urzúa, 2006). Under the context of racial wealth gaps, Barsky, Bound, Charles, and Lupton (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 the racial wealth gaps.

In my main analysis, I rely on the semi-parametric decomposition method introduced

[21] CHJP (2020) is mainly based on a regression of children's mean income *ranks* on their parents' income *ranks*, and they argue that this rank-rank relationship is actually close to linear for both Black and white men. It is unclear whether children's income (or income ranks) is also a linear function of individual skill, childhood neighborhood characteristics, or other family background characteristics.

by DiNardo, Fortin, and Lemieux (1996, hereafter DFL).[22] This method relaxes the parametric functional forms that the classical Oaxaca-Blinder decomposition imposes on the relationship between labor market outcomes (such as employment and earnings) and individual, family, and neighborhood characteristics. The DFL method also goes beyond analyzing racial gaps at the mean and examining racial gaps across the distribution of the outcomes. In a nutshell, this decomposition constructs the counterfactual distribution of labor market outcomes that I use to answer questions like, "What employment or earnings would white men in the NLSY–97 cohort have had if they had the same individual skill, family background, childhood neighborhood, or the school-to-work transition as Black men in the same cohort?" In the remaining part of this section, I describe the DFL method under the specific context of understanding racial labor market gaps.[23]

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, where the first line in Equation 2.1 below represents the racial gap that can be explained by Black-white differences in observed characteristics Z and the second line

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

[23] Readers familiar with the method may wish to skim through this section.

represents the unexplained residuals:

$$\begin{aligned} f_w(y) - f_b(y) &= f_w(y;Z_w) - f_w(y;Z_b) \\ &+ f_w(y;Z_b) - f_b(y;Z_b). \end{aligned} \tag{2.1}$$

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. 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 $Pr(z \mid b)$ 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. I check robustness of my decomposition results with different weight capping thresholds.[24]

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 emphasized in earlier studies (Barsky et al., 2002; Heywood and Parent, 2012). The intuition is straightforward: it is relatively less difficult to find white men at almost any point of the support of the Black distribution of Z, but it is sometimes quite difficult to find Black men at some parts of the white distribution of Z. Barsky et al. (2002) show that a lack of common support will likely introduce bias to the decomposition as more extrapolation is required. This common support problem will be exacerbated when the sample size is limited, which is especially relevant for the data of the NLSY–97.[25] I therefore stick to the decomposition in Equation 2.1 throughout my analysis.

In addition to the *aggregate* decomposition in Equation 2.1, the DFL method allows an estimation of the contribution of different subsets of variables in Z to the racial gap in labor

[24]Note that Altonji, Bharadwaj, and Lange (2012) also caps the weights in a similar context, where they apply the DFL decomposition and 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.

[25]Another distinction between the decomposition in Equation 2.1 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 the decomposition analysis. A main reason is that the earnings or employment function received by white men is arguably more similar to the hypothetical earnings or employment function in a labor market without discrimination (or other institutional barriers) against Black men.

market outcome y. This *detailed* decomposition helps answer questions such as, "What labor market outcomes would white men in the NLSY–97 cohort have achieved if they had the same family background and individual skill 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: childhood neighborhood N, family background F, individual skill S, and the school-to-work transition T. One of the possible detailed decompositions can be written as

$$\begin{aligned} f_w(y) - f_b(y) &= f_w(y; N_w, F_w, S_w, T_w) - f_w(y; N_b, F_w, S_w, T_w) \\ &+ f_w(y; N_b, F_w, S_w, T_w) - f_w(y; N_b, F_b, S_w, T_w) \\ &+ f_w(y; N_b, F_b, S_w, T_w) - f_w(y; N_b, F_b, S_b, T_w) \\ &+ f_w(y; N_b, F_b, S_b, T_w) - f_w(y; N_b, F_b, S_b, T_b) \\ &+ f_w(y; N_b, F_b, S_b, T_b) - f_b(y; N_b, F_b, S_b, T_b). \end{aligned} \tag{2.2}$$

The first line represents the contribution of Black-white differences in childhood neighborhood N. The contribution is the sum of a *direct* effect of childhood neighborhood N on labor market outcome y and an *indirect* effect, which comes from any changes in the distributions of F, S, and T that are attributed to the changes in N. In other words, this is the *unconditional* effect of neighborhood on the racial gap in y. The second line represents the contribution of Black-white differences in family background F *after* holding childhood neighborhood to be constant between Black and white men. It is important to note that when holding childhood neighborhood constant between Black and white men, any variations in family background that are implied by variations in childhood neighborhood 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 individual skill and

the school-to-work transition, respectively. The last line represents the racial gap in y that remains unexplained after accounting for Black-white differences in all observed factors in Z.

An important feature of the DFL decomposition is that the detailed decomposition is not unique. As is shown in Equation 2.2, the contributions of different components of Z to the overall racial gap depend on the sequential ordering by which the different components (N, F, S, and T) are added in to 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 orderings, but I always hold the relative positions of family background, individual skill, and the school-to-work transition as the following: Family background $\rightarrow$ Individual skill $\rightarrow$ School-to-work transition. The inclusion of family background before individual skill is justified by research demonstrating that family investments play a crucial role in the formation of skills (summarized in Cunha et al., 2006). I also always keep the school-to-work transition as the last component after all "pre-market" factors (including childhood neighborhood). I hold no prior as to where childhood neighborhood should be in the sequence relative to family background and individual skill, and I explore different positions of childhood neighborhood with the following orderings:

$$\text{Neighborhood} \rightarrow \text{Family} \rightarrow \text{Skill} \rightarrow \text{Transition}$$

$$\text{Family} \rightarrow \text{Neighborhood} \rightarrow \text{Skill} \rightarrow \text{Transition}$$

$$\text{Family} \rightarrow \text{Skill} \rightarrow \text{Neighborhood} \rightarrow \text{Transition}$$

In the first ordering, which is the same as Equation 2.2, childhood neighborhood is given the highest priority in the decomposition, while individual skill is given the lowest priority. The following two orderings move down the priority of childhood neighborhood in the sequence and move up family background and individual skill. In the second ordering, the contribution of childhood neighborhood is estimated as conditional on Black-white differences in family background. In the third ordering, the contribution of childhood neighborhood is estimated conditional on Black-white differences in both family background and individual skill.

Comparing the decomposition results *across* different orderings helps explain what the *unconditional* explanatory power of childhood neighborhood, as documented in CHJP (2020), actually represents. This is a direct examination of Heckman's comments (Heckman, 2018) on the interpretation of CHJP's findings. The difference between the estimated explanatory power of neighborhood in the first and the second ordering tells us the extent to which the unconditional neighborhood effect is an artifact of residential sorting of Black and white men with different *family background* into different neighborhoods.

The difference between the second and the third ordering further tell us about the relationship between childhood neighborhood and individual skill. In particular, if much of the estimated explanatory power of childhood neighborhood in the second ordering goes away as we move to the third ordering, it suggests either one (or both) of the following two stories. First, there may be residential sorting at the *individual* level that cannot be fully captured by the family variables in the NLSY–97, and the true effect of neighborhood in explaining racial labor market gaps is limited. Second, if there is a true effect of neighborhood, it is likely affecting racial gaps in labor market outcomes through the channel of influencing the skill accumulation process. On the other hand, if the estimated unconditional explanatory power of neighborhood (as in the first ordering) stays quantitatively robust as we move to the second and the third ordering, it shows that the neighborhood effect as

documented in CHJP (2020) is not simply reflecting residential sorting of individuals and families across neighborhoods (at least not sorting based on my included individual and family variables in the NLSY–97 data).

In all three orderings, I include the school-to-work transition as the last component because it is presumably partly the *outcome* of both pre-market factors and what happened to Black and white men when they left school. As I detail in the next section, I measure the school-to-work transition with one's weeks worked in the first year post-schooling. Conditional on individual skill, family background, and childhood neighborhood, the estimated contribution of the school-to-work transition to racial gaps observed at later stages (such as the sixth to eighth years post-schooling) can capture either the effect of Black-white differences in *unobserved* pre-market factors (such as certain non-cognitive skills that are difficult to measure) or other unobservables that are reflected in how Black and white men initiated their careers differently. An example of such an unobservable is discrimination against Black men both inside and outside the labor market (Donohue and Heckman, 1991; Pager, 2003; Bertrand and Mullainathan, 2004; Council of Economic Advisors, 2016). If Black men had a harder time in finding the first job due to discrimination, then at least part of the estimated contribution of transition T reflects the effect of discrimination. Another potential example is differential exposure to the Great Recession between Black and white men due to, for example, differences in residence at labor market entry and occupation and industry choices that *cannot* be explained by observed individual skill, family background, and childhood neighborhood.

The effect of any racial differences that cannot be fully captured by individual skill, family background, childhood neighborhood, *and* the school-to-work transition are left in the residuals. It is important to emphasize 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 (also known as "quantities"), and it does not reveal the potential effect of racial differences in

the returns paid to each one of these factors (also known as "prices"). Different prices paid to Black and white men will be absorbed in the residuals. It might be of particular interest to future decompose the residuals to see, for example, the specific contribution of racial differences in skill prices.[26]

The counterfactuals in Equation 2.2 can be estimated in a similar way as $f_w(y;Z_b)$ in Equation 2.1. For example, write $f_w(y;N_b,F_b,S_w,T_w)$, the counterfactual density of y for white men when they had the same childhood neighborhood and family background as Black men, as the following integral:

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

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

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

As explained earlier, $Pr(b \mid n,f)$ and $Pr(w \mid n,f)$ can be estimated with a probit model that includes N and F 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 in this specific case *and* in other cases with different orderings of N, F, S, and T.

[26]For example, by imposing a linear function form and focusing on the *mean*, the Oaxaca-Blinder decomposition can be used to quantify the fraction of the residuals that is driven by different prices paid to Black and white men for their N, F, S, and T. Firpo, Fortin, and Lemieux (2018) proposes a decomposition method based on re-centered influence function regressions that extends to general distributional statistics.

2.4 Description of Racial Differences in Individual Skill, Family Background, and Childhood Neighborhood

A decomposition of the observed racial labor market gaps into the contributions of individual skill, family background, childhood neighborhood, and the school-to-work transition requires a detailed and careful definition of these factors. In this section, I first describe which specific variables I include in each one of the four sets of factors and how I construct these variables from the NLSY–97 and from linking its geocode file to external sources.

The set of individual skills include four variables that the literature has shown to have an important impact on labor market outcomes. First, I control for formal schooling using the highest grade completed.[27] Second, I include the AFQT score, which has been extensively used as a measure of cognitive skills (Neal and Johnson, 1996; Heckman, Stixrud, and Urzúa, 2006; Urzúa, 2008; Lang and Manove, 2011).[28] The cohort were 12–18 years old when they took the test. I use the AFQT score constructed by the National Longitudinal Survey (NLS)'s team, which adjusts for different test-taking ages. In particular, I include dummy variables for the AFQT score deciles to allow for potential nonlinear effects.[29]

[27] I show robustness by controlling for schooling with dummy variables for educational groups: less than high school, high school graduate, some college, and college graduate and above.

[28] The AFQT score is created based on four sections of the Armed Services Vocational Aptitude Battery (ASVAB) test. An important question is whether the test is racially biased. If so, the measured racial gap in AFQT score is picking up biases in the test rather than racial differences in cognitive skills alone. A study led by the National Academy of Sciences (NAS) in 1991 examined the racial fairness of the test and concluded that the AFQT score does not systematically underpredict the job performance of Blacks relative to whites (Wigdor and Green Jr., 1991). It is largely an open question whether the result of this NAS study can be extended to civilian population, and there is a debate of this question in the literature (Rodgers and Spriggs, 1996; Heckman, 1998).

[29] I show robustness using two alternative cognitive skill measures. First, I include the age-adjusted AFQT score percentiles linearly. Second, I use the AFQT score constructed by Altonji, Bharadwaj, and Lange (2012), which also adjusts for different test-taking ages. The aim of Altonji, Bharadwaj, and Lange (2012) is to create comparable AFQT scores between the NLSY–97 and NLSY–79. For this purpose, the authors create a crosswalk of scores between the two cohorts, which requires stronger assumptions and external data sources. The AFQT score constructed by Altonji, Bharadwaj, and Lange (2012) is cardinal, while the score constructed by the NLS is ordinal. See Altonji, Bharadwaj, and Lange (2012) for more details about how they construct the score.

Last, it has been shown that non-cognitive skills are also important predictors for educational and labor market outcomes, even conditional on the effect of cognitive skills (Heckman, Stixrud, and Urzúa, 2006; Urzúa, 2008). A recent series of papers focuses on a specific type of non-cognitive skills, called social skills, showing that its importance seems to be especially high in today's labor market (Weinberger, 2014; Deming, 2017; Kahn and Deming, 2017). However, there is a lack of consensus about how to measure non-cognitive or social skills. In this chapter, I use the non-cognitive score and social score constructed by Deming (2017), based on personality trait questions in the NLSY–97.[30]

I measure family background with the following five variables. First, I control for annual parental income measured when children (i.e., the NLSY–97 respondents) were ages 12–16. Second, I control for the mother's education level with her highest grade completed. Third, I control for family structure with an indicator variable for living with both parents at age 14. Fourth, I include an indicator variable for whether the mother is a teenage mom, defined as being younger than 20 when giving birth to the child. Last, psychologists and sociologists have measured parenting style along two dimensions: strictness and supportiveness (Maccoby and Martin, 1983; Baumrind, 1991). Some recent work on parenting completed by economists also adopt this measure (Doepke and Zilibotti, 2017; Doepke, Sorrenti, and Zilibotti, 2019). In the NLSY–97, respondents are asked about how strict and supportive their mother is, and the answers are used to classify mothers into four groups following the literature: authoritative (strict, supportive), authoritarian (strict, not supportive), indulgent (not strict, supportive), or uninvolved (not strict, not supportive).[31]

[30] It is worth pointing out that the non-cognitive and social skill measures in the NLSY–97 are potentially subject to large measurement errors. Unlike the AFQT score, which is measured for the NLSY–97 cohort at presumably pre-market ages (12–18), the personality trait questions in the NLSY–97 were asked when the cohort were either 17–21 or 23–27. This means that the non-cognitive and social skill measures are possibly already influenced by one's labor market experiences and contain more (likely non-classical) measurement errors. As I present later, non-cognitive and social scores turn out to have minimal explanatory power to the racial labor market gaps in my data, and my decomposition results are quantitatively similar when excluding non-cognitive and social scores from the set of individual skill.

[31] Note that CHJP (2018) also consider basic family background information. They include parental income

The measure of childhood neighborhood quality includes two series of variables. The first series includes neighborhood quality measures at the county and higher levels, which I construct by using the geocode file of the NLSY–97 to link respondents' childhood county of residence at age 12 to external data sources. Chetty and Hendren (2018b) construct a county-by-county measure of neighborhood quality by comparing the income of children from families who move to "better" counties when children are younger with those who move when children are older.[32] Under certain assumptions, this measure can be thought of as a sufficient statistic for county neighborhood quality in terms of improving children's future income. The authors create this county-specific quality measure separately for men and women from high-income and low-income families. For the purpose of this chapter, I include a measure for men from high-income families and a measure for men from low-income families.[33]

In addition to the county quality measures, I also link childhood county of residence in the NLSY–97 to the 2000 Census to draw information on county socioeconomic conditions. Specifically, this information includes population, median household income, poverty rate, and the share of men with a college education. It is possible that at a more aggregate level, neighborhood quality has an independent effect other than the effect of county quality. For this reason, I also include commuting zone quality measures, similarly created by Chetty and Hendren (2018b), and state socioeconomic conditions from the 2000 Census.

The second series of childhood neighborhood measures aims to capture *within-county* neighborhood quality. First, I classify one's neighborhood at age 12 into five groups by

in their main analysis and also include the mother's education, family structure, and family wealth in some specifications. I show robustness using alternative sets of family background variables, such as excluding parental income.

[32]The county neighborhood quality measure is created by Chetty and Hendren (2018b) based on a sample of children born between 1980 and 1986, and their income is measured at age 26 (from federal income tax records).

[33]Low-income (high-income) families are families at the 25th (75th) percentile of the national family income distribution. The reason to include separate measures is that a "good" neighborhood for a certain group might not be a "good" neighborhood for another group.

whether it is in a metropolitan statistical area (MSA), in a central city, and in an urban or rural area.[34] The five groups are MSA central city, MSA non-central city urban area, MSA non-central city rural area, non-MSA urban area, and non-MSA rural area. Importantly, this classification captures the high density of Black families living in MSA central cities and the high density of white families living in MSA non-central city urban areas (i.e., suburban areas), which I document in the NLSY–97. Second, I include an indicator variable for whether the respondent lived in a house or apartment owned by the family when the respondent was 12–16. Other studies have documented that at the neighborhood-level, homeownership rate is positively associated with neighborhood quality and housing price (Coulson and Li, 2013), and at the family level, homeownership leads to more family investment in local amenities and social capital (DiPasquale and Glaeser, 1999). I therefore include homeownership as a proxy for neighborhood quality at a more local level.[35]

In theory, it is not obvious what is the most appropriate level of childhood neighborhood classification. Given the complexity of how neighborhood can influence children's outcomes, there may not be a simple answer to this question that applies to different children from different neighborhoods. In practice, due to the residential sorting of individuals and families, the more detailed neighborhood classification I choose (e.g., control for neighborhood quality at the census tract level instead of the county level), the more likely the neighborhood measures are capturing the characteristics of individuals and families living in the neighborhood rather than the characteristics of neighborhood itself. In the most extreme case, if neighborhood is classified at the dwelling level, then neighborhood and family characteristics will have an almost perfect overlap and become indistinguishable

[34] According to the Census Bureau, the urban-rural distinction is defined on a block-by-block basis, based on the population density (as well as some other characteristics) of each census block. So an MSA could include both urban and rural areas.

[35] It is debatable whether homeownership more closely captures family- or neighborhood-level characteristics. By assigning this variable to childhood neighborhood (rather than to family background), I give childhood neighborhood more priority when evaluating the contribution of different factors to the overall racial labor market gaps.

from each other.

In past studies, the choice of neighborhood classification usually depends on the specific research question and the data availability. For example, CHJP (2018) control for childhood neighborhood at the census tract or block level in their descriptive analysis, while they switch to the commuting zone level when estimating the causal effect of moving to a "better" neighborhood. As a comparison, my set of neighborhood quality measures *alone* explains 20%–30% of the observed average racial gap in annual earnings between Black and white men in the NLSY–97, while in CHJP's sample, census tract fixed effects *alone* explain 31% and census block fixed effects *alone* explain 44% of the observed average racial gap in income ranks.

The last component in my decomposition is the school-to-work transition. As shown in Figures 2.2–2.3 and in the top panel of Table 2.1, there are substantial racial gaps observed at various margins of employment outcomes in the first year post-schooling. To capture this, I measure the school-to-work transition flexibly with a vector of indicator variables for whether the respondent worked for 0 weeks, 1–9 weeks, 10–19 weeks, ..., 40–49 weeks, or 50 weeks and above.[36]

Table 3.1 summarizes the Black-white differences in individual skill, family background, and childhood neighborhood characteristics. Across most of the selected variables, there are large and statistically significant racial differences. For individual skill, Black men have lower education attainment, measured cognitive skills, and measured social skills. The Black-white difference in measured non-cognitive skills is indistinguishable from zero. In terms of the magnitude of these differences, the racial gap in measured cognitive skills is either about 27 percentiles, on average, (by the NLS's measure) or close to one standard

[36] A sufficient statistic for the school-to-work transition is the number of weeks before finding the first job, which is presented in Table 2.1. I do not use this measure because for some young men in my sample; they either did not start to work until toward the end of the eight-year period or never worked in the first eight years. For these men, the measure of actual transition duration has a time overlap with my outcome of interest, racial labor market gaps observed in the sixth to eighth year.

deviation (by Altonji et al.'s measure). As a comparison, the racial gap in measured social skills is about one-fifth standard deviation.

For family background, Black men in the NLSY–97 are less likely to grow up in a two-parent family, their parents have substantially lower income, and their mothers are less educated and are more likely to be teenage moms.[37] It is especially striking that only about 32% of Black men in the NLSY–97 lived with both parents at age 14, as compared to 63% of white men in the NLSY–97. Regarding parenting style, more than 80% of *both* Black and white mothers are reported by their child to be supportive. However, compared to white mothers, a larger share (about 15 percentage points more) of Black mothers are reported to be strict.

In terms of childhood neighborhood quality, Black men in the NLSY–97 are less likely to grow up in "good" counties and "good" commuting zones, according to the neighborhood quality measures constructed by Chetty and Hendren (2018b). Black men are also less likely to grow up in counties and states with higher median household incomes, lower poverty rates, and larger shares of college-educated men. In addition, substantially more Black men have grown up in central cities, while more white men have grown up in suburban areas (MSA, non-central city, urban areas). Seventy-five percent of white men lived in a house or apartment owned by their families around ages 12–16, while only 41% of Black men did. These patterns from Table 2 consistently suggest that there are large racial differences in the NLSY–97 regarding the quality of childhood neighborhoods.

Given the observed racial differences in individual skill, family background, childhood neighborhood, and the school-to-work transition, in the next section I evaluate how these factors have contributed to the overall racial gaps in employment and earnings, averaged over the sixth to eighth years post-schooling.

[37] If the respondent lives with a single parent, the parental income measure only includes the income of that parent. Part of the substantial parental income gap between Black and white men is due to a larger share of Black men living in single-parent families. I use the inverse hyperbolic sine to allow for zero income values.

2.5 Decomposition Results

2.5.1 Decomposing Racial Gaps at the Mean

How have individual skill, family background, childhood neighborhood, and the school-to-work transition contributed to the overall racial labor market gaps observed in the NLSY–97? In this section, I start by explaining racial gaps at the *mean* using the semi-parametric DFL decomposition. My main analysis uses the balanced sample of Black and white men who have not been enrolled in formal schooling for at least eight years. As previously explained, I focus on racial gaps in employment and earnings averaged over the six to eight years post-schooling, when the labor market outcomes have mostly stabilized.

Table 2.3 summarizes my main decomposition results. The three panels each feature a DFL decomposition that includes the same four sets of explanatory factors but with different *orderings*. As discussed extensively in Section 3, I hold the relative position of family background, individual skill, and the school-to-work transition fixed and alter the position of childhood neighborhood in the ordering. I also always keep the school-to-work transition as the last component because it is partly an outcome.[38] Within each panel, I first present the racial gap in employment (average weeks worked per year) and earnings (log of average annual earnings including zeroes), and then I present the share of the gap that can be explained by specific factors. The last column presents the share of the racial gap that is left unexplained and is in the residuals.

Two key findings stand out. The first is the central role of individual skill. No matter which sequential ordering is used, individual skill has the largest explanatory power to mean racial *earnings* gap. Measured racial differences in skills account for 42% of the

[38]The contribution of lower-order components is estimated after holding the higher-order components the same between Black and white men. Intuitively, this means that higher-order components are given some priority in claiming the explanatory power over the overall racial labor market gaps.

racial earnings gap when conditional on childhood neighborhood and family background, and 46% of the gap when only conditional on family background. This key role of individual skill is also observed in explaining the mean racial *employment* gap. Measured racial skill differences account for 36% of the racial employment gap when conditional on childhood neighborhood and family background and 43% of the gap when only conditional on family background. Note that family background shows a similar explanatory power as individual skill to the racial *employment* gap when family background is added as the first component in the sequential ordering.

Given the central role of racial skill differences, a natural question is whether a specific skill measure has driven this result. Recall that the set of skills includes highest grade completed, measured cognitive skills (AFQT score), and measured non-cognitive and social skills. Although it is difficult to disentangle the effects of different skill measures, as they can be endogenous to each other, I can explore, in a descriptive sense, which specific skill measure has the dominant explanatory power.[39]

In Table 2.4, I contrast decomposition results where I include only highest grade completed in the set of individual skill (middle panel) to results where I include only AFQT score as a measure of cognitive skills (bottom panel). As a benchmark, the top panel replicates the main results in Table 2.3. Comparing across panels in Table 2.4, it is clear that the explanatory power of individual skill is attributable primarily to measured cognitive skills (AFQT score) rather than formal schooling. Conditional on childhood neighborhood and family background, schooling accounts for only 3% of the racial employment gap and 4% of the racial earnings gap. The result barely changes when I add non-cognitive and social scores to schooling.[40] In stark contrast, conditional on childhood neighborhood and

[39]Urzúa (2008) shows that cognitive skills can grow as education attainment grows and people can make endogenous schooling decision based on their underlying cognitive and non-cognitive abilities.

[40]Recall that the measures of non-cognitive and social skills constructed in the NLSY–97 (Deming, 2017) are subject to potentially large measurement errors. In addition, the racial differences in measured non-cognitive and social skills are also not as sizable as the racial differences in measured cognitive skills, as

family background, AFQT score accounts for 31% of the racial employment gap and 36% of the racial earnings gap. The share accounted for by measured cognitive skills is very close to the share accounted for by the full set of individual skill (36% for employment, 42% for earnings). When only conditional on family background, the same pattern holds: the predominant contribution of skills is primarily driven by racial differences in measured cognitive skills.

When interpreting the central role of individual skill (especially cognitive skills), it is important to emphasize that the skill measures themselves shall be seen as an outcome. For example, cognitive skills in my data 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. Using a simple 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). In a cohort close in age to the NLSY–97, CHJP (2020) show 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 relative role of families and schools (or neighborhoods) 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). Identifying the specific mechanisms behind the racial skill differences in this cohort of young men is beyond the scope of this chapter and is left for future research.

The primary explanatory power of AFQT score, as compared to schooling, in explaining the racial labor market gaps in the NLSY–97, is consistent with our existing knowledge

shown in Table 3.1. It is therefore not surprising that adding non-cognitive and social scores or not barely changes the decomposition results.

based on previous cohorts of Americans. For example, using the data of the NLSY–79, Neal and Johnson (1996) show that AFQT score (in a quadratic function) alone accounts for about 60% of the wage gap between Black and white men, while schooling alone accounts for about 20%. My finding provides the first evidence for the Millennial cohort on the central role of cognitive skills in understanding the racial labor market gaps. It 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 tax records.[41]

This importance of incorporating a measure of cognitive skills leads to my second key finding in Table 2.3, which examines how the explanatory power of childhood neighborhood changes before and after accounting for racial differences in individual skill and family background that persist within neighborhoods. As discussed earlier, CHJP (2018) studies a cohort close in age to the NLSY–97 using tax records, but a significant limitation of their data is that they do not contain a direct measure of cognitive skills (such as the AFQT score).[42]

A vast body of evidence has shown that there is persistent residential segregation by race in the U.S., which has further limited the labor market prospects of Black Americans (e.g. Kain, 1968; Massey and Denton, 1993; Cutler and Glaeser, 1997; Charles, 2003). An important question about CHJP's finding is to what extent the documented neighborhood effects in their data reflect the residential sorting of Black and white families into different neighborhoods, a point raised by (Heckman, 2018).[43] The inclusion of rich individual and

[41] A potentially promising future research direction is to link tax records with survey datasets that have skill measures, such as the AFQT score in the NLSY–97 and the high school test scores in the National Education Longitudinal Study.

[42] Although the authors have education attainment in their data by linking tax records to the American Community Survey, my finding in Table 2.4 shows that it is really racial differences in measured cognitive skills, not differences in formal schooling, that have primary explanatory power.

[43] Since its release, the CHJP study has received tremendous attention from the press (e.g., *New York Times* 2018; *Washington Post* 2018) and has inspired constructive discussions among social scientists. Heckman (2018) raises a series of comments regarding both the findings of CHJP and what future research needs to

family variables (especially the skill measures) in the NLSY–97 dataset allows me to shed light on this question within the framework of the DFL decomposition.

In the top panel of Table 2.3, the contribution of childhood neighborhood is estimated *unconditionally*, as it is the first component in the ordering. In this case, my measures of childhood neighborhood characteristics explain 9% of the mean racial employment gap and 19% of the mean racial earnings gap. As I move childhood neighborhood down in the sequential ordering and estimate its contribution either conditional on family background (middle panel), or conditional on both family background and individual skill (bottom panel), the explanatory power of neighborhood to the racial labor market gaps diminishes substantially to small or zero. For the racial earnings gap, when conditional on family background, the contribution of neighborhood goes down from 19% to 11%. When further conditional on individual skill, the contribution of neighborhood further reduces to 7%. For the racial employment gap, when conditional on family background, the contribution of neighborhood goes down from 9% to –9%. The negative contribution of neighborhood means that holding family background the same between Black and white men *overcompensates* for Black disadvantage in childhood neighborhood characteristics. This finding is not totally surprising given that the *unconditional* contribution of neighborhoods to the racial employment gap is already low.[44]

This finding suggests that the observed unconditional effect of neighborhoods (as shown in the top panel), at least at the level that I can observe, can result from residential sorting of families and individuals. In sharp contrast, the estimated contribution of skills is gen-

address. In particular, he stresses the importance of reconciling different studies in the literature, which my findings help to do. See Heckman (2018) for more.

[44]The negative contribution of a specific factor in the DFL composition is not uncommon. For example, when studying how the labor market outcomes will change from the NLSY–79 to the NLSY–97, Altonji, Bharadwaj, and Lange (2012) find that the cross-cohort improvement in AFQT score has a negative contribution after conditioning on cross-cohort changes in family background characteristics. This is because the improvement in AFQT score is smaller than what would be predicted given the observed changes in family background across cohorts. In other words, the cross-cohort changes in family background overcompensate for the cross-cohort changes in AFQT score in their data.

erally robust whether conditional on childhood neighborhood or not (and it is always conditional on family background). Meanwhile, the estimated *unconditional* contribution of family background, which is 39% of the racial employment gap and 26% of the racial earnings gap, is also much greater than the estimated unconditional contribution of childhood neighborhood (9% for employment and 19% for earnings).

It is also worth pointing out that my finding here does not contradict the causal estimates of neighborhood effects in a series of recent studies (Chetty, Hendren, and Katz, 2016; Chyn, 2018; Chetty and Hendren, 2018a) but shows that the overall explanatory power of neighborhoods to the racial labor market gaps may be limited in this cohort. Basically, the causal neighborhood effects in these studies are estimated by comparing disadvantaged families (many of whom are Black families) who moved to "good" neighborhoods with disadvantaged families who stayed in their original neighborhoods or by comparing disadvantaged families who moved when children were younger with those who moved when children were older. This comparison, however, does not fully address the question of how the outcomes of the disadvantaged families who moved compare to families who were *already* living in the "good" neighborhoods. If, after moving, the Black families still fall substantially behind white families already there, it indicates that the neighborhood effects, though causal and significant, can actually have a limited power in explaining the overall racial income gaps, which is what my findings here suggest.

That said, my findings do not rule out the possibility that childhood neighborhood can have a true effect in explaining racial gaps in labor market outcomes. However, my finding of the diminishing explanatory power of measured childhood neighborhood characteristics *does* suggest that if there is a true neighborhood effect, it is mainly functioning through the channel of skill formation. As discussed previously, understanding where the racial skill gap comes from requires a formal investigation of the skill formation process, and the roles of families, schools, and neighborhoods in this process. This is beyond the scope of this

chapter.

In addition to the explanatory power of "pre-market" characteristics, for Black and white men in the NLSY–97, the difference in how they initiated their careers in the very first year post-schooling explains explains 13% of the racial employment gap and 14% of the racial earnings gap observed in the sixth to eighth years post-schooling. This contribution of the school-to-work transition is estimated *conditional on* racial differences in family background, individual skill, and childhood neighborhood. As discussed earlier, interpreting this result requires extra caution. Given the descriptive nature of my analysis, I do not attempt to tell how much of this estimated contribution of the school-to-work transition reflects the *causal* impact of one's initial work experience.[45]

My school-to-work transition measure should be thought as an *index* that could potentially capture three effects. First, it could capture racial differences in unobserved pre-market factors, such as non-cognitive skills that are difficult (if not impossible) to measure in nature and/or are measured with potentially nontrivial errors in the NLSY–97. Second, it could also pick up the effect of unobservables that are reflected in racial differences in initial work experience. For example, labor market discrimination against Black men, which I do not incorporate directly in my decomposition, could be the reason why Black men had a harder time finding the first job. The estimated contribution of the school-to-work transition thus captures the effect of labor market discrimination. Meanwhile, if Black men are more likely to be arrested due to racial discrimination in the criminal justice system (Council of Economic Advisors, 2016), and this lowers their chance to successfully initiate a career, then some of the estimated contribution of transition can also reflect non-labor

[45]The fundamental challenge in estimating the causal impact of one's initial work experience on later labor market performance is to distinguish heterogeneity from state dependence (Heckman, 1981). My analysis can be seen as an attempt to control for heterogeneity with observed characteristics in individual skill, family background, and childhood neighborhood. In ongoing analysis, I am linking measures of local labor market conditions (such as unemployment rates) at the time of school-exit to the NLSY geocode files. This will allow me to purge more exogenous variations in young men's initial work experiences.

market discrimination.[46] Third, my estimated contribution of transition could contain the causal impact of initial work experience on later labor market outcomes.

Comparison to the Oaxaca-Blinder Decomposition

After describing my main findings regarding racial labor market gaps at the mean, in this section I compare my semi-parametric DFL decomposition results to the linear Oaxaca-Blinder (hereafter OB) decomposition results. This comparison serves two purposes. First, some well-cited studies on either the previous cohort (Neal and Johnson, 1996) or this cohort (CHJP, 2018) are based on linear regressions, which are similar in principle to conducting an OB decomposition. Replicating my analysis using the OB decomposition will help better compare my findings to these studies. Second, by imposing stronger functional forms, I will be able to gain power and obtain more precise estimates.[47]

In Table 2.5, the top three panels are replicates of the DFL decomposition results from Table 2.3, and the bottom two panels are the OB decomposition results: one which includes all four sets of factors and one which includes only childhood neighborhood. The OB decomposition is based on linear regressions with all explanatory factors added together, and is order independent. For this reason, in an OB decomposition, the estimated contribution of neighborhood when it is estimated together with the other factors (second last panel) is not the same as when it is estimated on its own (last panel).[48]

[46] Other examples of such unobservables include racial differences in labor market networks and exposure to the Great Recession.

[47] Given the sample size of the NLSY–97, imposing more functional form restrictions to gain statistical power may be a reasonable decision to make. Note that the OB decomposition focuses on the *mean*. In my ongoing analysis, I am extending this exercise to racial gaps in general distributional statistics using the re-centered influence function method proposed by Firpo, Fortin, and Lemieux, 2018.

[48] To draw a direct comparison to my DFL decomposition, I use the regression coefficients estimated among white men for the OB decomposition. The key counterfactual estimated as in Table 2.5 for both the DFL and OB decomposition is therefore, "On average, what labor market outcomes would white men in the NLSY–97 have had if they had the same individual skill, family background, childhood neighborhood, and/or the school-to-work transition as Black men while holding the employment or earnings function unchanged?" As discussed in Section 3, due to the concern of overlapping support, I did not conduct the reverse decomposition

Regarding the relative contributions of individual skill, family background, and childhood neighborhood to the overall racial labor market gaps at the mean, the OB decomposition tells a *qualitatively* similar story. Racial differences in skills play a key role by explaining 20% of the mean racial employment gap and 30% of the mean racial earnings gap. Family background has a similar explanatory power as individual skill to the racial employment gap, but in terms of explaining the racial earnings gap, individual skill is the leading factor. When all four factors are added together in an OB decomposition (as the second last panel in Table 2.5 shows), the contribution of neighborhood is estimated conditional on racial differences in individual skill, family background, and the school-to-work transition. In this case, childhood neighborhood explains 7% of the racial employment gap and 9% of the racial earnings gap.

As a comparison, when added alone in an OB decomposition (as in the last panel), childhood neighborhood accounts for 24% of the racial employment gap and 22% of the racial earnings gap. Consistent with what we have seen in the DFL decomposition, unconditionally, childhood neighborhood explains a meaningful share of the racial labor market gaps, but this explanatory power diminishes substantially when conditional on racial differences that persist within neighborhoods. It is worth pointing out that the analysis in CHJP (2018) is mainly based on linear regressions, and it is therefore more comparable to an OB decomposition. Interestingly, under the OB decomposition, the unconditional explanatory power of my measures of neighborhood (which is 24% for employment and 22% for earnings) is larger than what it is under the DFL decomposition (which is 9% for employment and 19% for earnings). This shows that my choice of neighborhood measures can achieve an explanatory power in the NLSY–97 that is close to what census tract fixed effects can achieve in CHJP's sample.[49]

by assigning the characteristics of white men to Black men.

[49]Census tract fixed effects on its own explain about 31% of the observed racial income gap in the data of CHJP (2018).

Although the results are qualitatively similar between the DFL and the OB decompositions, there are *quantitative* differences that are worth discussing. The main difference is that in the OB decomposition, the estimated contribution of the school-to-work transition is larger than in the DFL decomposition. The school-to-work transition explains 32% of the racial employment gap and 22% of the racial earnings gap in the OB decomposition, while the shares explained by the school-to-work transition are 13% and 14%, respectively, in the DFL decomposition. At the same time, the shares explained by individual skill, family background, and childhood background are generally lower in the OB decomposition than in the DFL decomposition. This is partly because in the DFL decomposition, the effects of individual skill, family background, and childhood neighborhood are always estimated *unconditional* on the school-to-work transition (because transition is always the last component in the sequential ordering), while in the OB decomposition the effects are estimated *conditional* on the school-to-work transition.

Robustness Checks

Last in this section, I present two robustness checks on my main results in Table 2.3. Considering my key finding of the central role of cognitive skills, the first robustness check tests whether my finding is robust to the use of alternative measures of cognitive skills. Considering my choice of sample construction and the examination of the early career period, the second robustness check tests whether my finding is robust to alternative samples and alternative periods of analysis.

Table 2.6 presents the results of my first main robustness check, where I consider two alternative measures of cognitive skills. Recall that in my main analysis, I measure cognitive skills with dummy variables for the deciles of the AFQT score constructed by the NLS. The first alternative measure is a linear AFQT score percentile (which ranges from 1

to 100) constructed by the NLS. Unlike the decile dummies, this measure does not allow for the potential non-linear relationship between AFQT score and labor market outcomes. The second alternative measure is the AFQT score constructed by Altonji, Bharadwaj, and Lange (2012). As mentioned in Section 4, this measure is built on more assumptions. But its main advantage is that it is cardinal and potentially contains more information on racial skill differences than the ordinal measure constructed by the NLS. For example, in the sample of the NLSY–97, the AFQT score distribution is left-skewed. This means that the actual skill difference between the 1st and 9th percentile of the AFQT score distribution could be larger than the actual skill difference between the 91st and the 99th percentile.

The first panel in Table 2.6 replicates my main result, the middle panel presents the result using the linear AFQT percentile constructed by the NLS, and the bottom panel presents the result using the AFQT score constructed by Altonji, Bharadwaj, and Lange (2012). Looking across panels, the estimated contribution of individual skill is highly stable. When using the linear AFQT percentile, the estimated contribution is slightly lower, which is likely because it imposes a more restrictive functional form than the decile dummies. When using the AFQT score constructed by Altonji, Bharadwaj, and Lange (2012), the estimated contribution of skills is even higher. When conditional only on family background, it accounts for 47% of the racial employment gap and 51% of the racial earnings gap. As discussed above, one possible reason is that Altonji et al.'s score captures more information on racial skill gaps that are not contained in the NLS's measure. As a result of the increased explanatory power of individual skill, the explanatory power of the school-to-work transition goes slightly down (as it is estimated *conditional* on individual skill).

Table 2.7 presents my second main robustness check. Specifically, I first present decomposition results using labor market outcomes summarized over the second to eighth years post-schooling. In my analysis in Table 2.3, I focus on labor market outcomes summarized over the sixth to eighth years, when the employment and earnings trajectories in

the NLSY–97 had mostly stabilized. Looking at the second to eighth years can help us understand the explanatory power of different factors to the racial gaps observed through *full* early career trajectories.

The middle panel of Table 2.7 shows the result for this alternative outcome, and as a comparison, the top panel replicates the main result from Table 2.3.[50] The first noticeable difference between the two panels is that when looking at racial gaps through the full trajectories (i.e., the second to eighth years), the overall explanatory power of the four factors is lower, especially for the racial earnings gap. Recall from Figures 2.2–2.4 that one's labor market status fluctuates and experiences growth in the very first few years beyond schooling completion. It is therefore possible that compared to the more stable later stage (i.e., the sixth to eighth years), the full early career trajectories are influenced by more unobserved factors that were not included in my measures of individual skill, family background, childhood neighborhood, and the school-to-work transition.

Although the overall explanatory power of the four observed factors decreases, the primary role of individual skill still holds as previously shown. Across different orderings, racial differences in skills explain up to 38% of the racial employment gap and up to 39% of the racial earnings gap. Unconditionally, childhood neighborhood explains 15%–18% of the racial labor market gaps, but when conditional on family background and individual skill, its explanatory power substantially reduces to small or zero, which is also consistent with my main findings in Table 2.3.

In the bottom panel of Table 2.7, I present decomposition results using an unbalanced sample of Black and white men in the NLSY–97 who completed schooling at least two years previously, and I summarize labor market outcomes over the second to eighth years.

[50]Compared to labor market outcomes averaged over the sixth to eighth years, the racial employment gap averaged over the second to eighth years is larger (7.55 weeks per year versus 6.47 weeks per year), and the racial earnings gap is smaller (1.18 log points versus 1.45 log points). This is because throughout the early career years, Black men caught up relative to white men in the number of weeks worked but fell further behind in earnings.

As a comparison, my main analysis uses a balanced sample of young men who completed schooling at least eight years previously. This alternative sample is unbalanced because part of the sample has not reached the eighth year post-schooling, and when averaging employment or earnings over the second to eighth years, it is averaged over different numbers of years for different people.[51] Despite this drawback, there are two possible advantages of using this unbalanced sample. The first advantage is its larger sample size: the unbalanced sample includes 1,210 white men and 534 Black men, and the balanced sample (which I use for the main analysis) includes 796 white men and 367 Black men. Relatedly, the second advantage is that the unbalanced sample might be less subject to potential sample selection issues.

Here I compare the bottom panel with the middle panel of Table 2.7, as both panels focus on labor market outcomes averaged over the second to eighth years. The results using the unbalanced sample are generally consistent with the results using the balanced panel. Racial skill differences still play an important role in explaining racial labor market gaps, and the estimated contribution of childhood neighborhood still reduces to zero when conditional on family background and individual skill. The only meaningful difference is that family background turns out to have the largest explanatory power in magnitude. But it is important to keep in mind that the contribution of individual skill is always estimated conditional on family background. Given this, when individual skill is estimated only conditional on family background (and not on childhood neighborhood), its explanatory power (33% for employment and 35% for earnings) is smaller than, but close to, that of family background (45% for employment and 39% for earnings).

In results not presented here, I also conduct robustness checks regarding the choice of family background variables, the choice of measures of schooling, different ways to cap

[51] For example, some people in this unbalanced sample have only one year of observation, and some have eight years of observations. The pattern of how observations are distributed across early career years is similar between Black and white men.

the propensity score weights, and the use of earnings measures with imputation for missing values. In general, the main findings from Table 2.3 are stable across these variations.

2.5.2 Decomposing Racial Gaps across the Distribution

So far I have focused on understanding racial employment and earnings gaps at the mean. The average racial gap is an important and informative statistic, but in practice, public policies are sometimes tailored to serve more specific groups, such as the low-income population. In this section, I go beyond the mean and explore the contribution of individual skill, family background, childhood neighborhood, and the school-to-work transition to the racial gap observed across the employment or earnings distribution. This is facilitated by the DFL method, which estimates the whole counterfactual distribution of labor market outcomes in a semi-parametric manner. However, it is important to keep in mind throughout this section that the sample size of the NLSY–97, to some extent, limits my ability to examine racial gaps across the employment or earnings distribution in a precise way.

I start by visually presenting the actual and counterfactual distributions for Black and white men in the NLSY–97. Figure 2.5 plots the employment distributions, and Figure 2.6 plots the earnings distributions. In each panel of the two figures, I present two *actual* distributions (solid line for white men and long-dashed line for Black men) and a *counterfactual* distribution (short-dashed line). The counterfactual distribution in the upper panel of each figure is for white men if they had the same childhood neighborhood characteristics as Black men. The counterfactual distribution in the bottom panel of each figure is for white men if they had the same full set of individual skill, family background, childhood neighborhood, and the school-to-work transition as Black men.[52] The vertical lines repre-

[52]For the simplicity of display, I do not plot other possible counterfactuals, such as the counterfactual distribution for white men if they had the same childhood neighborhood *and* family background as Black men.

sent the *25th percentile* of the employment distributions for white men (solid line) and for Black men (long-dashed line).

Comparing the actual distributions of average weeks worked between Black (long-dashed line) and white men (solid line) in Figure 2.5 shows clearly that the racial employment gap is largely driven by fewer Black men working for close to full-year employment (50 or more weeks a year). As previously shown in Table 2.1, in the sixth to eighth years post-schooling, 60% of white men in my sample worked for at least 50 weeks per year, while only 38% of Black men did so.[53]

In the upper panel of Figure 2.5, when moving from the actual earnings distribution for white men to the counterfactual distribution when white men had the same childhood neighborhood as Black men (short-dashed line), the distribution shifts to the left but to a very limited degree. This suggests that only a small share of the large gap between Black and white employment distributions can be attributed to childhood neighborhood. The racial gap between the 25th percentile of the white employment distribution (solid vertical line) and the 25th percentile of the Black employment distribution (long-dashed vertical line) is about 15 weeks per year. If childhood neighborhood on its own explains a large share of this gap, we should expect the counterfactual (short-dashed vertical line) to be much lower than the actual 25th percentile of the white distribution (solid vertical line) and closer to that of the Black distribution (long-dashed vertical line). Figure 2.5 shows the opposite: the counterfactual is only slightly shifted to the left of the actual 25th percentile of the white distribution. Note that what I present visually here is consistent with the top panel in Table 2.3, where I focus on the mean.

The starkest comparison in Figure 2.5 is between the two counterfactual employment

[53] As 90% of white men and 78% of Black men worked for at least 26 weeks per year over the sixth to eighth years post-schooling, for the convenience of display, I only show the upper part of the employment distribution (26 weeks or more). The DFL decomposition is still conducted over the whole employment distribution.

distributions. Although holding childhood neighborhood the same between Black and white men only slightly narrows the racial gaps across the employment distribution, further incorporating family background, individual skill, and the school-to-work transition pushes the counterfactual employment distribution for white men (short-dashed line) leftward to a marked extent. In particular, there is no longer a dominating share of white men (compared to Black men) concentrated close to full-year employment. When focusing on the racial gap at the 25th percentile, it is clearly shown that accounting for racial differences in all four factors shifts the counterfactual (short-dashed vertical line) a long way toward the actual 25th percentile of the Black distribution (long-dashed vertical line). This relatively limited explanatory power of childhood neighborhood and the large overall explanatory power when incorporating family- and individual-level factors again are consistent with what was previously shown for the mean racial gap.

A similar visual pattern holds in Figure 2.6, where I plot the actual and counterfactual distributions for average earnings over the sixth to eighth years post-schooling.[54] In the upper panel, accounting for racial differences in childhood neighborhood shifts the counterfactual white earnings distribution (short-dashed line) only slightly. In the lower panel, where I further account for racial differences in family background, individual skill, and the school-to-work transition, the counterfactual white earnings distribution is substantially closer to the *actual* Black employment distribution (long-dashed line). This is consistent with what Table 2.3 indicates for racial earnings gap at the mean.

The visual presentation in Figures 2.5–2.6 intuitively shows the overall explanatory power of family background, individual skill, and the school-to-work transition over childhood neighborhood on the racial gaps across the employment and earnings distribution.

[54]There is a mass of zeroes at the left end of the earnings distribution for both Black and white men (and the mass is larger for Black men). As there are almost no observations for a wide range of values above zeroes, for the convenience of display, I only show the upper part of the earnings distribution (log earnings above eight). The DFL decomposition is still conducted over the whole earnings distribution.

The next step is to explore in more detail what each of these factors has contributed at different parts of the distribution. Tables 2.8–2.9 summarize the *detailed* decomposition results for employment and earnings, respectively. As mentioned above, considering my sample size, from here on I focus my attention on racial gaps at the 25th percentile, the median, and the 75th percentile of the distribution. Since the racial employment gap is zero at the 75th percentile (as more than a quarter of both Black and white men on average worked for a full year over the sixth to eighth years post-schooling), I do not including the decomposition result for this particular outcome.

The first thing to note is that, for both employment (Table 2.8) and earnings (Table 2.9), the racial gap at the 25th percentile is significantly larger than racial gaps at the median or the 75th percentile. For example, young men at the 25th percentile of the *Black* employment distribution worked 15 fewer weeks per year than young men at the 25th percentile of the *white* employment distribution. The gap is about 6 weeks per year at the median. A similar pattern is observed for the racial earnings gaps.

I first focus on employment in Table 2.8. Two key findings previously observed at the mean (as in Table 2.3) are consistently observed for at the 25th percentile. First, individual skill plays a central role by explaining up to 52% of the racial employment gap. Second, the estimated contribution of childhood neighborhood is zero when conditional on family background and individual skill. Similarly to what I have previously shown at the mean, the negative contribution of childhood neighborhood (as long as it is estimated conditional on family background) is because holding family background the same between Black and white men overcompensates for racial differences in childhood neighborhood. In addition, observed racial differences in the four factors together account for 82% of the racial employment gap. A distinctive result at the 25th percentile is that the school-to-work transition now has a greater explanatory power. Conditional on the three pre-market factors, the school-to-work transition explains 25% of the racial employment gap at the 25th percentile,

as compared to 13% at the mean (Table 2.3).

As noted above, the racial employment gap at the median is much smaller than the gap at the 25th percentile. Table 2.8 shows that the observed factors' overall explanatory power to the gap at the median is also substantially lower. The four factors together account for about 40% of the gap, much of which is attributable to the contribution of the school-to-work transition. Notably, when only conditional on family background, individual skill explains 28% of the racial employment gap, which is much larger than the explanatory power of family background or childhood neighborhood. Comparing between the decomposition results at the 25th percentile and the median of the employment distribution, it appears that both individual skill and family background are making a greater contribution to racial gap at the lower part of the employment distribution.

I then turn to earnings. Table 2.9 presents the decomposition results for the racial earnings gap at the 25th percentile, the median, and the 75th percentile. The two key findings at the mean are again consistently observed at different parts of the earnings distribution. First, individual skill turns out to be the primary contributor, and it explains up to 29% of the racial gap at the 25th percentile and up to 38% of the racial gap at the 75th percentile. Second, although childhood neighborhood shows a meaningful explanatory power unconditionally, its explanatory power diminishes to zero when conditional on family background and individual skill.

When comparing the contribution of a specific factor across different parts of the earnings distribution, it appears that both family background and individual skill are making a balanced contribution to racial gaps at the lower and upper part of the earnings distribution. In contrast, the school-to-work transition shows a substantially larger explanatory power to the racial gap at the lower part of the earnings distribution. At the 25th percentile, 31% of the racial earnings gap can be attributed to the school-to-work transition. At the median, 20% of the gap can be attributed to the school-to-work transition. At the 75th percentile,

the explanatory power of transition is even smaller. What does this quantitatively large role of transition for more disadvantaged young men imply? It suggests that helping disadvantaged Black men to get a foothold in the labor market, which might be facilitated through innovative training programs, is a potentially important pathway to reduce racial labor market gaps at later career stages.[55]

To summarize, my key findings from Table 2.3, the central role of skills and the small or zero contribution of neighborhoods once conditional on family background and individual skill, are similarly observed when I conduct the decomposition across the employment and earnings distribution. In addition to what Table 2.3 shows for the mean, individual skill and family background appear to have a greater explanatory power to racial gap at the lower part of the *employment* distribution, and the school-to-work transition appears to have a greater explanatory power to racial gap at the lower part of the *earnings* distribution.

2.6 Conclusion

Millennials are playing an increasingly important role in various aspects of economic, social, and political life in the United States, but our knowledge to date of this cohort is still limited. In this chapter, I analyze the early careers of this cohort of young men, with a special focus on the racial gaps between Black and white men. I show a substantial racial gap in employment and earnings that largely persists through the first eight years beyond schooling completion. The data I use, the NLSY–97 and its restricted geocode file, allows me to conduct a coherent decomposition analysis to explore how individual-, family-, and neighborhood-level factors have contributed to the racial labor market gaps.

[55] Although the effect of traditional government training programs has been shown to be modest on average and limited among youths (see a review by Friedlander, Greenberg, and Robins, 1997), there are recent examples of job training programs designed and led by non-government organizations that show encouraging results for helping disadvantaged youths initiate a career. One example is Year Up, which involves potential employers in the training program (Fein and Hamadyk, 2018).

My key finding is that racial skill differences play a central role in explaining the racial gaps in employment and earnings observed in the early careers of Millennial men. This is primarily driven by racial differences in measured cognitive skills rather than by differences in formal schooling. Additionally, I find that conditional on racial differences in family background and individual skill, the explanatory power of my measures of childhood neighborhood characteristics is small or zero. I also show that the observed racial differences in the factors that I control for explain the vast majority of the racial employment and earnings gaps in the NLSY–97. 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 polices.

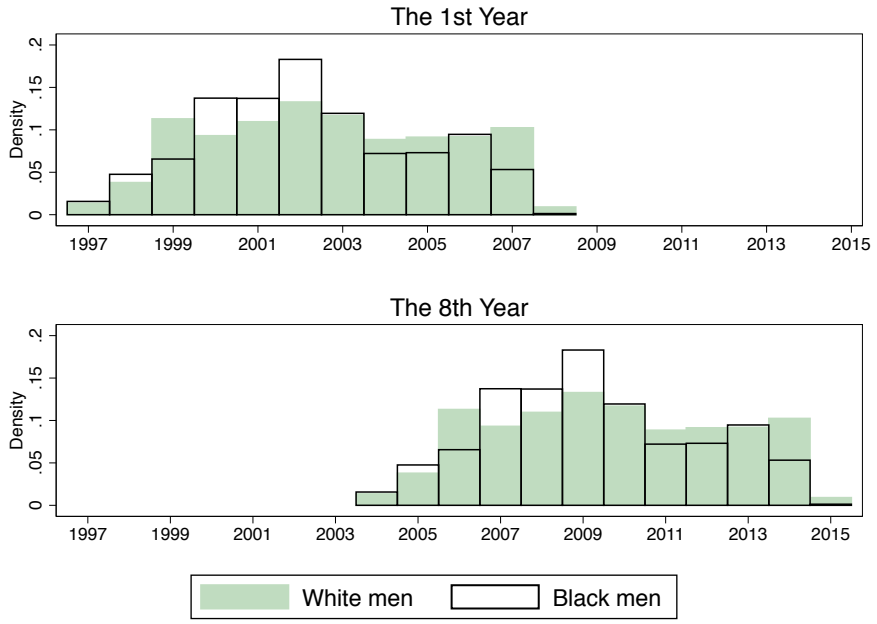
Despite the dramatic changes in both the characteristics of young men and the overall structure of the U.S. labor market, cognitive skills turn out to be the key driver of racial labor market gaps among Millennial men, as in previous cohorts. This suggests that although market demand for skills might have evolved over the past few decades, cognitive skills are still rewarded in today's labor market, and are particular drivers of racial gaps.[56]

Meanwhile, while the *average* skill level among young Americans has gone up across cohorts (Altonji, Bharadwaj, and Lange, 2012), evidence shows that racial skill gaps, by various measures, have either stayed stagnant or grown larger from 1980s to early 2000s (Neal, 2006). The stagnant or growing racial skill gaps, together with my finding of the importance of skills in explaining the racial labor market gaps among this cohort of men, strongly suggest that attention needs to be paid to understanding the skill accumulation process and more importantly, Black disadvantage in this process. Specifically, potentially effective pathways to reduce racial labor market gaps include public programs that foster

[56]Recent studies have documented declines in the returns to cognitive skills across cohorts (Castex and Dechter, 2014; Deming, 2017). Hellerstein, Luo, and Urzúa (2019) show that this finding is dependent on how the distribution of cognitive skills has changed across cohorts. Once holding the skill distribution fixed at the level of the previous cohort, the estimated returns to cognitive skills do not go down.

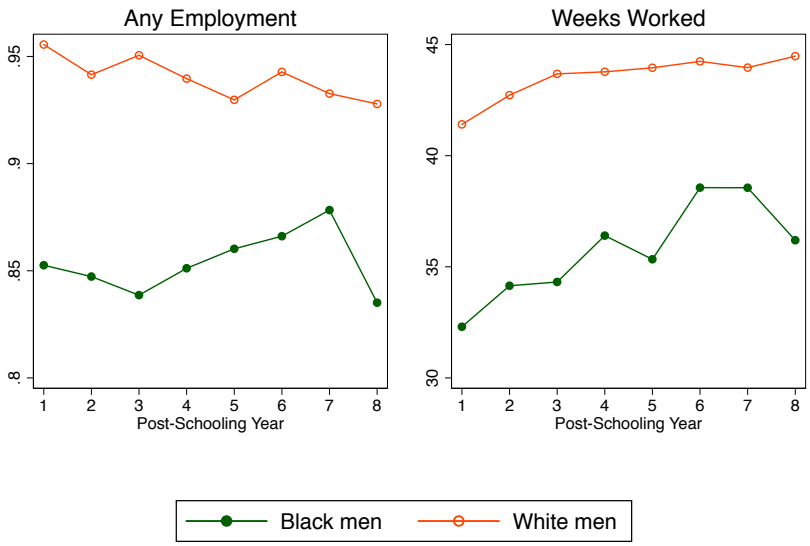
skill accumulation among Black men. For example, existing evidence based on previous cohorts suggests that family investments in young children have especially high returns (Cunha et al., 2006). Identifying the mechanisms behind the racial skill differences among Millennials will be important for designing policies for this cohort and beyond.

Figure 2.1: Corresponding Calendar Years for the First and Eighth Year Post-Schooling



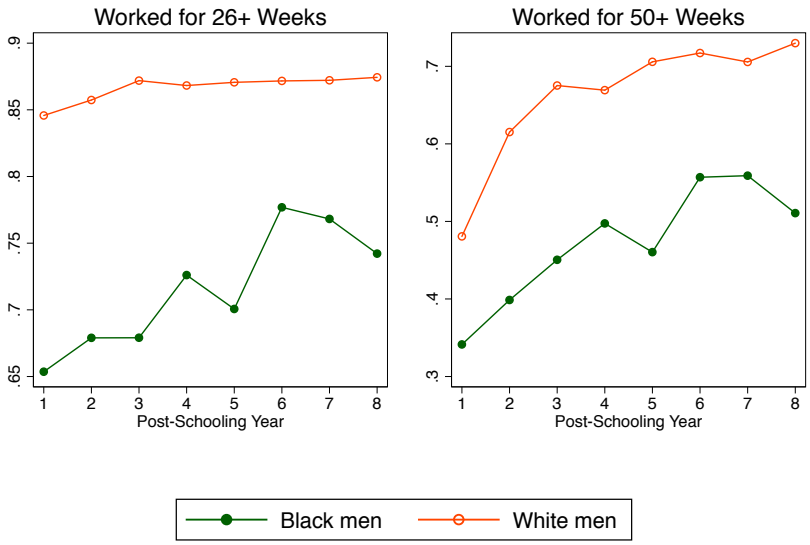
Notes: The upper panel specifies what the corresponding calendar years are for the first year post-schooling. The lower panel specifies the corresponding calendar years for the eighth year post-schooling. The sample is a balanced panel of young men who have completed schooling for eight years. Sample weights are used.

Figure 2.2: Employment Trajectories of Black and White Men in the NLSY–97



Notes: The left panel shows any employment in a year, and the right panel shows the number of weeks worked in a year. Any employment is defined as working for at least one week.

Figure 2.3: Employment Trajectories of Black and White Men in the NLSY–97 (continued)

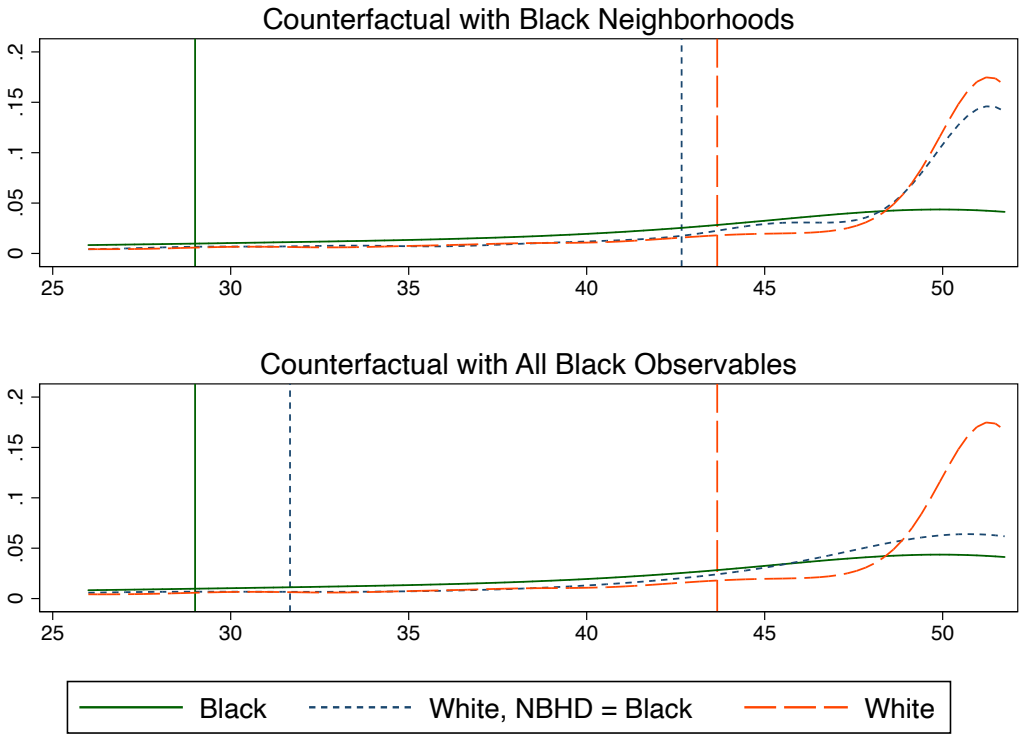


Notes: The left panel shows employment for at least 26 weeks in a year (i.e., half year), and the right panel shows employment for at least 50 weeks in a year (i.e., full year).

Figure 2.4: Annual Earnings Trajectories of Black and White Men in the NLSY–97

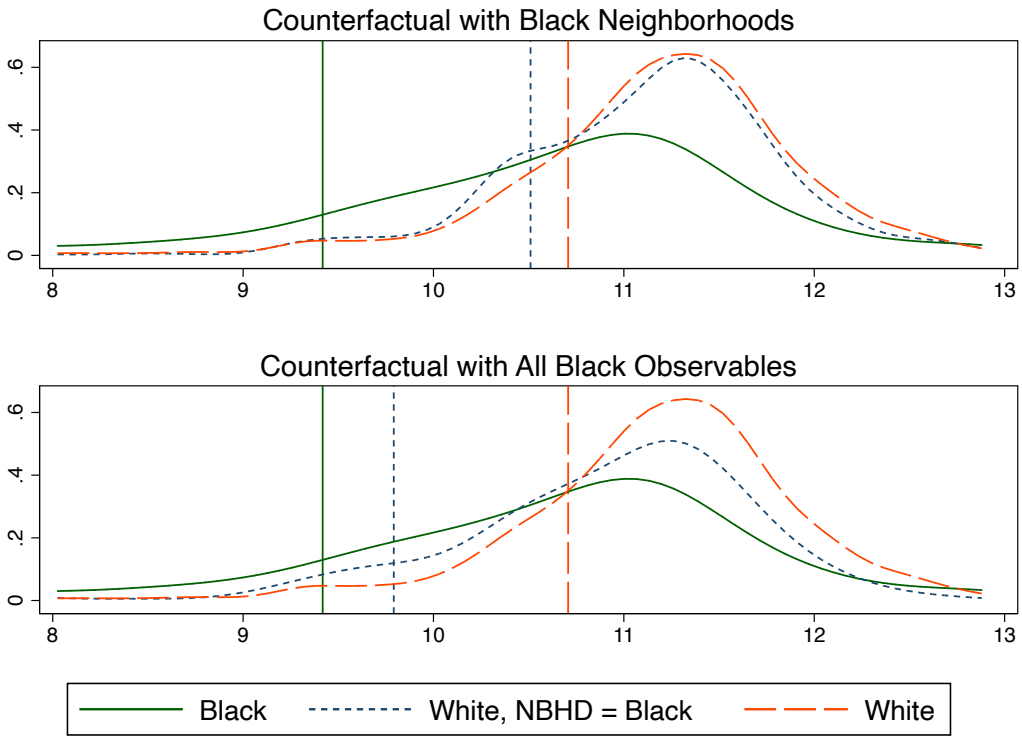
Notes: The left panel shows annual earnings, and the right panel shows the inverse hyperbolic sine of annual earnings, which I label as "log earnings" for simplicity. Earnings are adjusted to 2013 dollars.

Figure 2.5: Actual and Counterfactual Employment Distribution



Notes: The solid line represents the employment (average weeks worked per year) distribution for white men, and the long-dashed line represents the employment distribution for Black men. The short-dashed line represents the counterfactual employment distribution for white men if they had the same characteristics as Black men. The upper panel shows the counterfactual with Black childhood neighborhood, and the lower panel shows the counterfactual with Black observables in all four sets of factors. The vertical lines are the corresponding 25th percentile of the actual and counterfactual distributions. For the convenience of display, I only show the upper part of the employment distribution. The decomposition is still conducted over the whole employment distribution.

Figure 2.6: Actual and Counterfactual Earnings Distribution



Notes: The solid line represents the earnings distribution for white men, and the long-dashed line represents the earnings distribution for Black men. The short-dashed line represents the counterfactual earnings distribution for white men if they had the same characteristics as Black men. The upper panel shows the counterfactual with Black childhood neighborhood, and the lower panel shows the counterfactual with Black observables in all four sets of factors. The vertical lines are the corresponding 25th percentile of the actual and counterfactual distributions. For the convenience of display, I only include log earnings above eight in this figure. The decomposition is still conducted over the whole distribution.

Table 2.1: Early Career Experiences of Black and White Men in the NLSY–97

	White	Black	White-Black Gap	*p*-value
Initial Stage (the 1st Year)				
Weeks before the first job	10.53	31.83	–21.29	0.00
Any employment	0.96	0.85	0.10	0.00
Worked for $\geq$ 26 weeks	0.85	0.65	0.19	0.00
Worked for $\geq$ 50 weeks	0.48	0.34	0.14	0.00
Weeks worked	41.41	32.30	9.11	0.00
Log annual earnings (excl. zeroes)	10.40	9.94	0.46	0.00
Log annual earnings	9.43	7.28	2.15	0.00
Annual earnings (excl. zeroes)	23,305	18,786	4,520	0.02
Annual earnings	21,142	13,752	7,391	0.00
Later Stage (Averaging the 6th–8th Years)				
Weeks worked per year	44.18	37.72	6.47	0.00
Worked for $\geq$ 26 weeks per year	0.90	0.78	0.12	0.00
Worked for $\geq$ 50 weeks year	0.60	0.38	0.21	0.00
Log average annual earnings (excl. zeroes)	11.19	10.67	0.52	0.00
Log average annual earnings	10.57	9.12	1.45	0.00
Average annual earnings (excl. zeroes)	45,108	32,643	12,466	0.00
Average annual earnings	41,664	27,068	14,597	0.00
Summarizing the Early Career (the 1st–8th Years)				
Number of non-employment (NE) spells	1.74	2.67	–0.93	0.00
Average duration of NE spells (months)	6.61	10.05	–3.45	0.00
Cumulative weeks worked	346.84	285.11	61.73	0.00
Log average annual earnings (excl. zeroes)	10.99	10.44	0.56	0.00
Log average annual earnings	10.74	9.57	1.17	0.00
Average annual earnings (excl. zeroes)	36,827	24,484	12,343	0.00
Average annual earnings	34,548	20,566	13,982	0.00

[1] The sample is a balanced panel of 796 white men and 367 Black men who have completed schooling for at least eight years. Sample weights are used.

[2] Earnings are adjusted to 2013 dollars. Inverse hyperbolic sine is used to include zero values.

Table 2.2: Descriptive Characteristics of Black and White Men in the NLSY–97

	White	Black	White-Black Gap	P-value
Individual Skill				
Highest grade completed	13.17	12.08	1.09	0.00
AFQT percentile (NLS)	52.88	26.19	26.69	0.00
AFQT score (Altonji et al. 2012)	169.85	139.72	30.13	0.00
Non-cognitive score	–0.12	–0.06	–0.06	0.46
Social score	–0.03	–0.23	0.20	0.01
Family Background				
Log parental income at ages 12–16	11.04	9.07	1.96	0.00
Mother's highest grade completed	13.07	12.54	0.53	0.00
Living with both parents at age 14	0.63	0.32	0.31	0.00
Mother is not a teenage mom	0.86	0.74	0.12	0.00
Mother's parenting style				
Strict, supportive	0.41	0.55	–0.14	0.00
Strict, not supportive	0.09	0.10	–0.01	0.60
Supportive, not strict	0.40	0.29	0.11	0.01

[1] The sample is a balanced panel that includes 796 white men and 367 Black men who have completed formal schooling for at least eight years. Sample weights are used. MSA stands for metropolitan statistical area.

(Continued) Descriptive Characteristics of Black and White Men in the NLSY–97

Childhood Neighborhood	White	Black	White-Black Gap	P-value
Living in house or apartment owned by family	0.75	0.41	0.34	0.00
Residence type				
MSA, central city	0.20	0.36	–0.16	0.00
MSA, non-central city, urban area	0.34	0.23	0.11	0.00
MSA, non-central city, rural area	0.21	0.19	0.01	0.68
Non-MSA, rural area	0.17	0.13	0.04	0.24
Neighborhood Quality (Chetty and Hendren 2018b)				
County quality for low-income families	0.07	–0.29	0.36	0.00
County quality for high-income families	0.05	–0.05	0.09	0.00
Commuting zone quality for low-income families	0.06	–0.27	0.33	0.00
Commuting zone quality for high-income families	0.02	–0.10	0.11	0.00
County Socioeconomic Conditions				
Log population	12.20	12.43	–0.23	0.06
Log median household income	10.97	10.89	0.08	0.00
Poverty rate	0.11	0.15	–0.04	0.00
Male college rate	0.24	0.23	0.01	0.19
State Socioeconomic Conditions				
Log population	15.80	15.78	0.02	0.74
Log median household income	10.98	10.95	0.04	0.00
Poverty rate	0.12	0.13	–0.01	0.00
Male college rate	0.26	0.25	0.01	0.00

[1] The sample is a balanced panel that includes 796 white men and 367 Black men who have completed formal schooling for at least eight years. Sample weights are used. MSA stands for metropolitan statistical area.

Table 2.3: Contribution of Individual, Family, and Neighborhood Factors

	Racial Gap	Share Explained by				Residuals
DFL Ordering		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	36%	13%	21%
Log avg annual earnings	1.45	19%	18%	42%	14%	7%
DFL Ordering		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	36%	13%	21%
Log avg annual earnings	1.45	26%	11%	42%	14%	7%
DFL Ordering		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	43%	–16%	13%	21%
Log avg annual earnings	1.45	26%	46%	7%	14%	7%

[1] NBHD stands for neighborhood; DFL stands for the DiNardo-Fortin-Lemieux method. The three panels use different orderings of family background, individual skill, childhood neighborhood, and the school-to-work transition.

[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 are in the inverse hyperbolic sine. The sample includes 796 white men and 367 Black men who have completed formal schooling for at least eight years. Sample weights are used.

Table 2.4: Explanatory Power of Schooling and Cognitive Skills

	Racial Gap	Share Explained by				Residuals
Full Skill Set		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	36%	13%	21%
Log avg annual earnings	1.45	19%	18%	42%	14%	7%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	36%	13%	21%
Log avg annual earnings	1.45	26%	11%	42%	14%	7%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	43%	–16%	13%	21%
Log avg annual earnings	1.45	26%	46%	7%	14%	7%
Only Formal Schooling		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	3%	13%	54%
Log avg annual earnings	1.45	19%	18%	4%	12%	47%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	3%	13%	54%
Log avg annual earnings	1.45	26%	11%	4%	12%	47%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	5%	–11%	13%	54%
Log avg annual earnings	1.45	26%	4%	10%	12%	47%
Only Measured Cognitive Skills		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	31%	8%	30%
Log avg annual earnings	1.45	19%	18%	36%	10%	17%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	31%	8%	30%
Log avg annual earnings	1.45	26%	11%	36%	10%	17%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	41%	–18%	8%	30%
Log avg annual earnings	1.45	26%	40%	7%	10%	17%

[1] NBHD stands for neighborhood. The top panel uses the full skill set, which includes highest grade completed, cognitive (AFQT) score in decile dummies, and non-cognitive and social scores. The middle panel includes only schooling (highest grade completed) in the skill set. The bottom panel includes only cognitive (AFQT) scores in decile dummies in the skill set. The sample includes 796 white men and 367 Black men and is consistent across panels. 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 are in the inverse hyperbolic sine.

Table 2.5: Comparison of DFL and Oaxaca-Blinder Decomposition

	Racial Gap	Share Explained by				Residuals
DFL Ordering		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	36%	13%	21%
Log avg annual earnings	1.45	19%	18%	42%	14%	7%
DFL Ordering		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	36%	13%	21%
Log avg annual earnings	1.45	26%	11%	42%	14%	7%
DFL Ordering		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	43%	–16%	13%	21%
Log avg annual earnings	1.45	26%	46%	7%	14%	7%
Oaxaca-Blinder		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	7%	22%	20%	32%	19%
Log avg annual earnings	1.45	9%	17%	30%	22%	22%
Oaxaca-Blinder		NBHD				
Avg weeks worked per year	6.47	24%				76%
Log avg annual earnings	1.45	22%				78%

[1] NBHD stands for neighborhood; DFL stands for the DiNardo-Fortin-Lemieux method. The top three panels are replicated from Table 2.3. The bottom two panels are results from classical OB decomposition, one with all four factors and one with only childhood neighborhood.

[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 are in the inverse hyperbolic sine.

[3] The sample includes 796 white men and 367 Black men who have completed formal schooling for at least eight years. Sample weights are used.

Table 2.6: Robustness of Different Measures of Cognitive Skills

	Racial Gap	Share Explained by				Residuals
AFQT decile dummies (NLS)		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	36%	13%	21%
Log avg annual earnings	1.45	19%	18%	42%	14%	7%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	36%	13%	21%
Log avg annual earnings	1.45	26%	11%	42%	14%	7%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	43%	–16%	13%	21%
Log avg annual earnings	1.45	26%	46%	7%	14%	7%
AFQT percentile (NLS)		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	34%	16%	20%
Log avg annual earnings	1.45	19%	18%	37%	16%	10%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	34%	16%	20%
Log avg annual earnings	1.45	26%	11%	37%	16%	10%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	42%	–17%	16%	20%
Log avg annual earnings	1.45	26%	42%	6%	16%	10%
AFQT score (Altonji et al. 2012)		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	42%	11%	16%
Log avg annual earnings	1.45	19%	18%	47%	11%	5%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	42%	11%	16%
Log avg annual earnings	1.45	26%	11%	47%	11%	5%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	47%	–13%	11%	16%
Log avg annual earnings	1.45	26%	51%	7%	11%	5%

[1] NBHD stands for neighborhood. The top panel uses dummies for AFQT deciles constructed by the National Longitudinal Survey (NLS) team. The middle panel uses the (linear) AFQT percentile constructed by the NLS. The bottom panel uses the AFQT score constructed by Altonji et al. (2012). The sample includes 796 white men and 367 Black men. 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 are in the inverse hyperbolic sine.

Table 2.7: Robustness of Different Outcomes and Different Samples

	Racial Gap	Share Explained by				Residuals
Balanced Sample: 6th–8th Years		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.47	9%	21%	36%	13%	21%
Log avg annual earnings	1.45	19%	18%	42%	14%	7%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.47	39%	–9%	36%	13%	21%
Log avg annual earnings	1.45	26%	11%	42%	14%	7%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.47	39%	43%	–16%	13%	21%
Log avg annual earnings	1.45	26%	46%	7%	14%	7%
Balanced Sample: 2nd–8th Years		NBHD	Family	Skill	Transition	
Avg weeks worked per year	7.55	15%	18%	21%	18%	27%
Log avg annual earnings	1.18	18%	13%	29%	8%	32%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	7.55	32%	1%	21%	18%	27%
Log avg annual earnings	1.18	23%	8%	29%	8%	32%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	7.55	32%	39%	–17%	18%	27%
Log avg annual earnings	1.18	23%	38%	–1%	8%	32%
Unbalanced Sample: 2nd–8th Years		NBHD	Family	Skill	Transition	
Avg weeks worked per year	6.26	24%	22%	16%	17%	21%
Log avg annual earnings	1.02	23%	18%	20%	10%	29%
		Family	NBHD	Skill	Transition	
Avg weeks worked per year	6.26	45%	1%	16%	17%	21%
Log avg annual earnings	1.02	39%	2%	20%	10%	29%
		Family	Skill	NBHD	Transition	
Avg weeks worked per year	6.26	45%	33%	–17%	17%	21%
Log avg annual earnings	1.02	39%	35%	–13%	10%	29%

[1] NBHD stands for neighborhood. The top panel uses the balanced sample and focuses on outcomes averaged over the sixth to eighth years. The middle panel keeps the balanced sample but focuses on outcomes averaged over the second to eighth years. The bottom panel uses the unbalanced sample that includes men who have completed schooling for at least two years and focuses on outcomes averaged over the second to eighth years.

[2] The balanced sample includes 796 white men and 367 Black men. The unbalanced sample includes 1,210 white men and 534 Black men. Different sample weights are used for the two samples.

Table 2.8: Contribution to Racial Gaps across the Employment Distribution

	Racial Gap	Share Explained by				Residuals
DFL Ordering		NBHD	Family	Skill	Transition	
25th percentile	14.67	7%	18%	32%	25%	18%
Median	6.00	0%	6%	6%	28%	61%
DFL Ordering		Family	NBHD	Skill	Transition	
25th percentile	14.67	34%	–9%	32%	25%	18%
Median	6.00	11%	–6%	6%	28%	61%
DFL Ordering		Family	Skill	NBHD	Transition	
25th percentile	14.67	34%	52%	–30%	25%	18%
Median	6.00	11%	28%	–28%	28%	61%

[1] NBHD stands for neighborhood; DFL stands for the DiNardo-Fortin-Lemieux method. The three panels include all four sets of factors with different orderings. The decomposition is conducted for racial employment gaps at the 25th percentile and at the median. The sample includes 796 white men and 367 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.

Table 2.9: Contribution to Racial Gaps across the Earnings Distribution

	Racial Gap	Share Explained by				Residuals
DFL Ordering		NBHD	Family	Skill	Transition	
25th percentile	1.29	15%	8%	18%	31%	28%
Median	0.67	9%	7%	14%	20%	51%
75th percentile	0.41	21%	10%	32%	0%	36%
DFL Ordering		Family	NBHD	Skill	Transition	
25th percentile	1.29	23%	1%	18%	31%	28%
Median	0.67	20%	–5%	14%	20%	51%
75th percentile	0.41	33%	–2%	32%	0%	36%
DFL Ordering		Family	Skill	NBHD	Transition	
25th percentile	1.29	23%	29%	–10%	31%	28%
Median	0.67	20%	20%	–11%	20%	51%
75th percentile	0.41	33%	38%	–8%	0%	36%

[1] NBHD stands for neighborhood; DFL stands for the DiNardo-Fortin-Lemieux method. The three panels include all four sets of factors with different orderings. The decomposition is conducted for racial employment gaps at the 25th percentile, the median, and the 75th percentile. The sample includes 796 white men and 367 Black men who have completed formal schooling for at least eight years. Sample weights are used.

[2] Annual earnings are averaged over the sixth to eighth years and are in the inverse hyperbolic sine.

Chapter 3: Have Early Career Racial Gaps Changed Across Two Cohorts of American Men?

3.1 Introduction

The second chapter of my dissertation examines the observed racial labor market gaps in the cohort of Millennials using the data of the National Longitudinal Survey of Youth 1997 (NLSY–97). In this chapter, I extend the analysis and make a cross-cohort comparison of the racial labor market gaps observed in the early careers of Millennials and Baby Boomers. In particular, I ask how the underlying forces behind racial labor market gaps have changed across the two cohorts of young American men. This analysis is facilitated by two similarly constructed NLSY datasets, the NLSY–79 and the NLSY–97, which are nationally representative samples of young Americans born in 1957–1964 and 1980–1984, respectively.[1]

I track the two cohorts of young men from early adulthood to their mid-30s and document the racial gaps in their early career employment and earnings trajectories. The schooling and work history files are similarly constructed for the NLSY–79 and the NLSY–97 cohorts, facilitating a valid cross-cohort comparison. Methodologically, I follow the second chapter and extend it to the NLSY–79 cohort. In my main analysis, I focus on racial gaps in weeks worked per year and annual earnings observed over the sixth to eighth years

[1] Throughout this chapter, I refer to the NLSY–79 as the "older" cohort and the NLSY–97 as the "younger" cohort.

post-schooling, when the labor market outcomes of young men in the two cohorts reach a relatively stable stage. Following the second chapter, I also examine what underlying factors explain the documented racial labor market gaps in the two cohorts by applying the semi-parametric decomposition method first introduced by DiNardo, Fortin, and Lemieux (1996).

In the decomposition, I mainly examine three pre-market factors that the literature emphasizes as important determinants of racial gaps in labor market outcomes: skills and education, family background, and childhood neighborhood characteristics.[2] Harnessing the richness of the NLSY datasets and following the literature, I measure these factors in as detailed a way as possible. To keep the decomposition results comparable between the two cohorts, in my main analysis I follow Altonji, Bharadwaj, and Lange (2012) and include only the individual, family, and neighborhood variables that are similarly constructed or can be appropriately concorded.[3] I also examine the role of the school-to-work transition, conditioning on racial differences in these pre-market factors.

The first question of interest is how much of the observed racial employment and earnings gaps can be explained by pre-market racial differences in education and skills, family background, and childhood neighborhood, measured at as detailed a level as the data allow.

[2]Much of the literature on racial gaps in the U.S. labor market uses data of older cohorts of Americans, such as the NLSY–79. This literature is enormous, and the discussed explanations range from pre-market factors, such as skills and education (Neal and Johnson, 1996; Heckman, Stixrud, and Urzúa, 2006; Urzúa, 2008), family background (Thompson, 2018), and childhood neighborhood (Chetty et al., 2020) to the school-to-work transition (Schwandt and Wachter, 2019) and discrimination (Donohue and Heckman, 1991; Pager, 2003; Bertrand and Mullainathan, 2004; Council of Economic Advisors, 2016).

[3]Some variables included in the decomposition analysis in the second chapter of this dissertation are not available for the NLSY–79. The variables included in this chapter are therefore a subtest of the variables included in the second chapter. In this chapter, the set of individual education and skills variables is the same as the second chapter, including highest grade completed, AFQT score (as a measure of cognitive skills), non-cognitive and social test scores. The set of family variables here includes family income, family structure, and mother's education, while the second chapter includes two extra variables: mother's parenting style, and whether mother is a teenage mom. The set of neighborhood variables here includes a series of socioeconomic conditions of county and state of childhood residence, and whether an individual lives in central cities, suburban areas, or rural areas. In some of my analysis, I also present the results when the extra variables are included in the NLSY–97. I discuss variable definitions in details in the next section.

In the NLSY–79 cohort, the decomposition shows that around 80% of the observed racial gaps in employment and earnings can be explained by measured racial differences in these three pre-market factors. In the NLSY–97 cohort, a smaller but still important share, about 50%, of the racial employment and earnings gaps can be explained by measured racial differences in the three pre-market factors of the younger cohort. The Oaxaca-Blinder decompositions reveal a similar pattern.

Why is a greater share of the racial labor market gaps in the NLSY–97 cohort unexplained than in the NLSY–79 cohort? The decomposition results imply that the early career experiences of this younger cohort are more subject to the influences of unobservable factors. I present suggestive explanations of what these unobservable factors might be.[4] First, some pre-market factors can be hard to measure, and the variables I use may not be sufficiently detailed. Focusing solely on the NLSY–97 data, which includes a richer set of variables measuring family background and childhood neighborhood, I show that accounting for racial differences in additional family and neighborhood variables does further explain some of the observed racial labor market gaps in the NLSY–97 cohort. In particular, the unconditional explanatory power of measured childhood neighborhood characteristics increases from 1%–9% to 10%–15% when I add more neighborhood variables that are only available in the NLSY–97 data. This finding is qualitatively consistent with the empirical findings of Chetty et al. (2020) and the second chapter of this dissertation.

Second, Black men in both NLSY cohorts initiated their careers with significantly worse employment and earnings outcomes (in the first year post-schooling) compared to their white counterparts, and past studies have shown that poor outcomes early in the

[4]One important unobservable factor, which I did not explore directly in this chapter, is racial discrimination against Black men in the labor market. There are indirect ways through which labor market discrimination affects my decomposition results. For example, due to discrimination, Black men may receive lower labor market returns to the same skills they own. Any racial differences in labor market returns will be left in unexplained residuals of the decomposition. I leave it for future research to formally investigate how racial discrimination has evolved overtime in the U.S. labor market.

school-to-work transition has a long-lasting impact on later labor market outcomes (Kahn, 2010; Schwandt and Wachter, 2019). Some of the unexplained racial labor market gaps could be due to Black young men not having a successful school-to-work transition, and the importance of a successful transition could be larger in the younger cohort. Conditioning on racial differences in the aforementioned pre-market characteristics, I further show that in the NLSY–97 cohort, racial differences in the school-to-work transition, measured as the number of weeks worked in the first year post-schooling, explain 10% of the racial gaps in employment and earnings observed six to eight years post-schooling.[5] In contrast, this conditional explanatory power of the school-to-work transition is not found in the NLSY–79.

In addition to the overall explanatory power of pre-market characteristics, a second question of interest is whether any specific factor is of particular importance in explaining the observed racial labor market gaps. In their influential paper, Neal and Johnson (1996) use the NLSY–79 data and show that simply controlling for racial gaps in cognitive skills, as measured by the Armed Forces Qualification Test (AFQT) score, eliminates 60% of the racial wage gaps between Black and white men.[6] This result has greatly shaped our understanding to date of the central role of skills in explaining racial gaps in the U.S. labor market. My DFL decomposition confirms this finding and shows that in the NLSY–79 cohort, racial differences in skills and education *alone* explain 60%–70% of the racial gaps in employment and earnings observed in the sixth to eighth year post-schooling. I also

[5]When further including county unemployment rate (UR) in the year of labor market entry as an additional measure of the school-to-work transition, the total explanatory power of transition reaches 16% of the racial employment and earnings gaps. The data of county UR are available only for the NLSY–97 sample years.

[6]Urzúa (2008) takes a more structural approach and makes an important distinction between measured cognitive skills (AFQT score) and underlying cognitive ability. The model in Urzúa (2008) specifies the observed (AFQT) test score as a function of both underlying (cognitive) ability and family background (e.g. family income or parental education). The author shows that cognitive ability explains about 40% of the racial gaps in wages and earnings in the NLSY–79, which is somewhat smaller than the explanatory power of AFQT score documented in Neal and Johnson (1996). This is because the racial gap in the AFQT score, according to the model, also picks up racial differences in family background, which have a direct effect on labor market outcomes.

show that in the NLSY–97 cohort, the share of the racial employment and earnings gaps explained by education and skills *alone* is about 30%. In both cohorts, the explanatory power of education and skills is primarily driven by measured racial gaps in cognitive skills (AFQT score) rather than by highest grade completed or measures of social and non-cognitive skills.

Why has the explanatory power of education and skills decreased across cohorts? One reason is that from the NLSY–79 to the NLSY–97 cohort, racial gaps in cognitive skills (measured by AFQT score) have fallen to a great extent, although the gaps in the younger cohort remain quantitatively substantial and statistically significant.[7] However, even though the overall explanatory power of observed pre-market factors is lower in the younger cohort, education and skills still remain the central observable explanatory factor of the racial labor market gaps.[8]

The last question I explore is where the racial gaps in education and skills come from and whether the mechanisms have changed across cohorts. I explore the relationship between education and skills with family and neighborhood characteristics by exploiting the sequentially nature of the DFL decomposition. Specifically, in both cohorts, I compare the *unconditional* explanatory power of racial differences in education and skills with the explanatory power after conditioning on measured racial differences in family background and childhood neighborhood characteristics.

In the NLSY–79 cohort, the unconditional explanatory power of education and skills goes away almost entirely after conditioning on family and neighborhood characteristics.

[7] Another plausible reason is changing returns to skills. If the returns to cognitive skills have declined across cohorts, the share of the racial labor market gaps explained by racial skills gaps could decline as well. Imposing a linear functional form and assuming a single skill price, past research shows that the wage returns to cognitive skills have declined from the NLSY–79 to the NLSY–97 cohort (Castex and Dechter, 2014; Deming, 2017). However, when these restrictions are relaxed, Hellerstein, Luo, and Urzúa (2019) shows that the returns to cognitive skills have not declined across cohorts.

[8] This is consistent with the findings of the second chapter of this dissertation, which focus on the NLSY–97 cohort.

Conversely, in the NLSY–97 cohort, the explanatory power of education and skills decreases only a modest amount and largely remains. This suggests that in the older cohort, the documented crucial role of education and skills in explaining racial labor market gaps mainly comes from racial differences in the family and neighborhood characteristics that the NLSY data allow me to measure, while in the younger cohort, factors beyond these family and neighborhood characteristics play a greater role.[9]

The remainder of this chapter proceeds as follows. In Section 2, I describe the NLSY datasets and how I create the concordance of variables between the two cohorts. I then show the racial differences observed in pre-market characteristics, including skills and education, family background, and childhood neighborhood, and present the early career labor market trajectories for Black and white men in both cohorts. Section 3 discusses the semi-parametric decomposition results. Section 4 concludes.

3.2 Comparing Two Cohorts of Young American Men

3.2.1 Data: NLSY–79 and NLSY–97

The main datasets I use throughout this chapter are 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 1957–1964 and 1980–1984, respectively. My analysis uses Black and white men from both the main sample and the minority subsample.[10]

The NLSY dataset fits the purpose of my analysis in three important ways. First, it includes a monthly diary of school enrollment and a weekly diary of work status, which

[9] I use the age-adjusted AFQT score constructed by Altonji, Bharadwaj, and Lange (2012). Part of the differences in the decomposition results between the two NLSY cohorts, especially those concerning the role of skills, could be due to potential imperfections of the score adjustment process.

[10] I do not include the economically disadvantaged white subsample or the military subsample of the NLSY–79.

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. I define schooling completion following the literature (Light and McGarry, 1998; Neumark, 2002).[11] In my main analysis, I keep a balanced panel of young men who completed schooling at least eight years (or 96 months) prior and track their labor market outcomes through the first eight years post-schooling.

Second, the NLSY records rich information on individual, family, and neighborhood characteristics, which is of critical importance to my decomposition analysis. In particular, it includes a measure of cognitive skills (AFQT score) that has been shown by past studies as a key determinant in understanding racial gaps in the U.S. labor market (Neal and Johnson, 1996).

Third, and most importantly, the NLSY–79 and the NLSY–97 surveys are designed and administrated in a similar way that much of the key variables from the two cohorts are comparable either directly or after some concordance, facilitating a valid comparison between the two cohorts. 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 the restricted-use geocode files.

3.2.2 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 diary 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 diary started, the exact school-exit time cannot be identified. To minimize this issue without losing too much sample size, I

[11] 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.

therefore exclude NLSY–79 respondents who were older than 18 as of 1980.[12] For both cohorts, I also exclude young men who were already out of school when the enrollment diary started or were still enrolled in school as of the most "recent" wave.[13]

Second, as of the most recent wave of the NLSY–97 cohort in 2015, the respondents were around ages 30–34. I focus my analysis of the NLSY–79 cohort to survey years 1979–1996, so in the most "recent" wave of 1996, the NLSY–79 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 adjusts for different test-taking ages and test format changes between the two cohorts, and I use their adjusted AFQT score.

Unlike the AFQT score for cognitive skills, there is no consistent measure of non-cognitive or social skills in 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 and convenient way to handle this incomparability 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.

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

[13] In other words, my sample includes young men who completed schooling after the enrollment diary started and before the most "recent" wave.

My set of family background includes 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 childhood.[14] In some of my empirical analysis, I include two more family variables that are only available in the NLSY–97 data: whether the respondent's mother is a teenage mom and the mother's parenting style (Doepke and Zilibotti, 2017).

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, and for the NLSY–79 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.[15] 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 a urban or rural area.[16] In some cases, I also include neighborhood variables that are only available in the NLSY–97 data: homeownership status at the first survey and a set of neighborhood quality measures at the county and commuting zone levels created by Chetty and Hendren (2018b).[17]

[14]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.

[15]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 create major incomparability 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 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.

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

[17]Chetty and Hendren (2018b) creates neighborhood quality measures separately for men and women from high- and low-income families. I use the two measures created for men from high-income families and men from low-income families.

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.

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

How have the racial gaps in pre-market education and skills, family background, and childhood neighborhood characteristics changed across cohorts? Altonji, Bharadwaj, and Lange (2012) create an index of skills for young Americans of the two NLSY cohorts and shows that the racial skill gap has fallen, on average, between Black and white men from the NLSY–79 cohort to the NLSY–97 cohort.[18] In Table 3.1, I compare Black and white men in each cohort along the series of pre-market characteristics discussed in the previous section and test whether the racial gap in each factor has changed significantly across cohorts.

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 score 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. This finding is consistent with the findings of Altonji, Bharadwaj, and Lange (2012), which are largely driven by the cross-cohort change in AFQT score. The racial gap in social test score has increased significantly, but the magnitude of change

[18]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. See Altonji, Bharadwaj, and Lange (2012) for details.

is arguably modest (about 0.35 standard deviations). The racial gap increases slightly in highest grade completed and decreases slightly in non-cognitive test score. 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 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 from counties and states with a larger population, lower median household income, higher poverty rate, and lower share of men with college education. Some of these racial gaps in neighborhood socioeconomic conditions appear to have fallen from the NLSY–79 cohort to the NLSY–97 cohort (such as county median household income and poverty rate). During their childhood, Black men are more likely to live in central cities and white men are likely to live in suburban areas (MSA, non-central city, urban areas). This racial difference seems to have also decreased across cohorts.

3.2.4 Racial Gaps in Early Career Trajectories

In this section I show how the racial gaps in early career trajectories have evolved across the two NLSY cohorts. I focus on employment and earnings outcomes in the first eight years after a young man completed schooling. In Table 3.2, I summarize the early career outcomes in three periods: the transition stage, defined as the first year post-schooling; the later stage, defined as the sixth to eighth years; and the entire first eight 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.

In both cohorts, Black men fell substantially behind their white counterparts in the transition stage along multiple margins of employment and earnings. It took Black men 30 more weeks to get the first job in the NLSY–79 cohort, and it took Black men 22 more weeks in the NLSY–97 cohort. The decrease in the racial gap is quantitatively meaningful but is 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). Black men worked for 13 fewer weeks in the first year in the NLSY–79 cohort and 9 fewer weeks in the NLSY–97 cohort. The racial gaps in all of these employment outcomes have fallen across cohorts, but only some of the declines can be distinguished from zero. The racial gap in annual earnings are large and significant in both cohorts, but the change from the NLSY–79 cohort to the NLSY–97 cohort is minimal.

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, and the increase was greater for Black 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 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, but the cross-cohort difference is not statistically distinguishable from zero.

For a large share of young men in the NLSY–97 sample, their early career years overlapped with the Great Recession, as shown in Figure 3.4. Because Black men have been

documented to suffer more from economic downturns than their white counterparts (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 employment outcomes in their first few years out of school.

Although the racial labor market gaps at two snapshots (the first year and the sixth to eighth years) have not changed by a statistically detectable amount, the *shapes* of employment and earnings trajectories could have exhibited more apparent changes. Figures 3.1–3.3 plot various employment and earnings outcomes year by year through the first eight complete years post-schooling. The starkest pattern is that employment and earnings trajectories have flattened from the NLSY–79 cohort to the NLSY–97 cohort.

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. In the NLSY–97 cohort, the employment and earnings outcomes of both races either stayed largely stable through 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 the Millennial cohort 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 3.1–3.3 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. In the decomposition analysis in the next section, I primarily focus on racial gaps in employment and earnings outcomes measured over the sixth to eighth years post-schooling (as shown in the middle panel of Table 3.2).

Last, the bottom panel in Table 3.2 summarizes racial labor market gaps over the full first eight years post-schooling. In both cohorts, Black men experienced more and longer

non-employment spells, worked fewer weeks cumulatively, and earned more. From the NLSY–79 cohort to the NLSY–97 cohort, the racial gaps in employment outcomes have fallen, while the racial gap in earnings have increased. The cross-cohort changes summarized over full eight years are not statistically significant.

3.3 Decomposition Results

In my main analysis, I apply the semi-parametric decomposition method introduced by DiNardo, Fortin, and Lemieux (1996, hereafter DFL). The second chapter of my dissertation has a detailed discussion of the method, and I refer readers who are interested in methodological details to the second chapter and/or the relevant section in Altonji, Bharadwaj, and Lange (2012), who also applies the DFL method.[19]

The DFL decomposition shows how much of the racial employment and earnings gaps, which I measure 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 I mainly focus on three pre-market characteristics: education and skills, family background, and childhood neighborhood characteristics. In some cases I also examine the explanatory power of racial differences in the school-to-work transition, which I measure as weeks worked in the first year post-schooling, and I always estimate the explanatory power of transition after conditioning on racial differences in the three pre-market characteristics.

3.3.1 Overall Explanatory Power of Observable Pre-Market Factors

My first empirical finding concerns how much of the observed racial labor market gaps can be attributed to pre-market racial differences in education and skills, family back-

[19]Following Altonji, Bharadwaj, and Lange (2012), I restrict the influence of outliers by imposing a cap to the DFL propensity weights.

ground, and childhood neighborhood *altogether*. Table 3.3 summarizes the result. Column 1 presents the raw racial gaps in employment (weeks worked) and earnings, as in the middle panel of Table 3.2. Column 2 presents the share of the raw racial gap that are explained by measured racial differences in the three pre-market factors together, and column 3 presents the share that remains unexplained.

In the NLSY–79 cohort, the DFL decomposition shows that 86% of the racial employment gap and 77% of the racial earnings gap are explained by racial differences in education and skills, family background, and childhood neighborhood. This means that the vast majority of the racial labor market gaps observed sixth to eighth years post-schooling in the older cohort can be attributed to Black disadvantages at the individual, family, and neighborhood levels observed before labor market entry. In the NLSY–97 cohort, the overall explanatory power of the *same* set of pre-market characteristics has decreased to 48% of the racial employment gap and 55% of the racial earnings gap.

The classical Oaxaca-Blinder decomposition, which imposes a restricted linear functional form, shows a similar pattern. In the NLSY–79 cohort, 93%–95% of the racial employment and earnings gaps can be explained by measured racial differences in the three pre-market characteristics, and this share has decreased to 40%–55% in the NLSY–97 cohort.

The reduction in the overall explanatory power of observed pre-market characteristics is consistent with a broad convergence between Black and white men in these characteristics. As shown in Table 3.1, the Black-white gaps have fallen along various individual, family, and neighborhood variables, including measured cognitive skills (AFQT score), mother's education, childhood residence type, and childhood county/state median household income and poverty rate. Under the assumption that the returns to the pre-market characteristics are constant across cohorts, a convergence in these characteristics will mechanically lead to a lower overall explanatory power of these characteristics in the younger cohort. Note

that the racial gaps in social test score and parental income have increased across cohorts. But the decomposition result indicates that the divergence in these two variables has been dominated by the convergence in other variables in terms of explaining racial labor market gaps in the two cohorts.

As the racial gaps in observed pre-market characteristics have converged across cohorts, the racial employment and earnings gaps in the sixth to eighth years post-schooling have not fallen accordingly. This suggests that in the younger cohort, racial differences in *unobservable* factors have played a more important role in shaping racial gaps in early career labor market outcomes.[20] What could such unobservable factors be? I discuss three potential possibilities in the following subsection.

Exploring Hard-to-Measure Variables Using the NLSY–97 Data

One reason for unobservable factors is there can be racial differences in pre-market characteristics that are rewarded in the labor market but are hard to measure. If these characteristics have become more important in the younger cohort than in the older cohort, a larger share of the observed racial labor market gaps in the younger cohort will be left unexplained in my decomposition analysis. For example, a series of recent studies have argued that the effect of childhood neighborhood on adulthood outcomes happens at a very local level, such as census tract or census block (Chetty, Hendren, and Katz, 2016; Chetty and Hendren, 2018a; Chetty et al., 2020). It is possible that the neighborhood measures in my main decomposition analysis (as in Table 3.3) are not detailed enough to capture the full effect of racial differences in neighborhood characteristics.

[20] In principal, another potential reason why a larger share of the racial gaps in the NLSY–97 cohort are left unexplained is model misspecification. However, the DFL decomposition imposes relatively little parametric restriction on the relationship between labor market outcomes and underlying characteristics, and there is no obvious reason to believe why model misspecification is a larger issue in the NLSY–97 cohort than in the NLSY–79 cohort. Also, the consistent results between the DFL decomposition and the Oaxaca-Blinder decomposition suggest that model specification is not a main concern in this context.

Luckily, when solely focusing on the NLSY–97 cohort, I am able to incorporate a richer set of neighborhood variables that are available in the NLSY–97 cohort but not in the NLSY–79 cohort. As described earlier, these variables include homeownership status in 1997 and a set of neighborhood quality measures at the county and commuting zone levels (Chetty and Hendren, 2018b). Examining how adding these variables affects the decomposition results in the NLSY–97 cohort will help me understand how much a concern the measurement issue is. Additionally, the NLSY–97 cohort also includes a few more family background variables that are not available in the NLSY–79 cohort: whether mother is a teenage mom and the mother's parenting style.

In Table 3.4, the top two panels replicate the main result of Table 3.3, and the bottom panel presents the result for the NLSY–97 cohort when the extra family and neighborhood variables are included. Two important patterns are revealed. First, when childhood neighborhood characteristics are added *alone* in the decomposition (and neither education and skills nor family background variables are added), the explanatory power goes up by a meaningful degree when the extra variables in the NLSY–97 cohort are included. This is shown in column 6 of Table 3.4.

When the extra variables are not included in the NLSY–97 cohort, measured racial differences in childhood neighborhood *alone* explains 1% of the racial employment gap and 9% of the racial earnings gap. After adding the extra neighborhood variables in the NLSY–97 cohort, the explanatory power goes up to 11% for employment and 15% for earnings. The Oaxaca-Blinder decomposition shows a similar pattern. In contrast, the extra family variables do not add explanatory power to the existing family variables as much as the extra neighborhood variables do to the existing neighborhood variables. This finding means that the extra neighborhood variables in the NLSY–97 cohort do capture some of the unobserved racial differences in childhood neighborhood, while the extra family variables do not.

In addition to the sole explanatory power of family or neighborhood variables, the *overall* explanatory power of all pre-market characteristics in the DFL decomposition stays almost unchanged with the extra family and neighborhood variables. Under the Oaxaca-Blinder decomposition, the overall explanatory power increases by a very limited degree, from 55% to 62% for employment and from 49% to 57% for earnings. This lack of change in the overall explanatory power of all pre-market factors is because much of the racial differences in these extra neighborhood variables are already reflected in the racial differences in the existing education and skills and family background variables. For example, some of the explanatory power of racial differences in childhood neighborhood characteristics, when added alone, is capturing the underlying characteristics of individuals and families who live in different neighborhoods rather than the effect of neighborhoods per se. When the decomposition already accounts for racial differences in education and skills and family background, it is not surprising that adding extra neighborhood variables does not increase the overall explanatory power much.

The main takeaway of Table 3.4 is that the extra neighborhood variables in the NLSY–97 cohort do seem to capture some of the unobserved racial differences at the neighborhood level, but adding these variables does not increase the overall explanatory power of pre-market by a meaningful degree. To explain the large unexplained racial employment and earnings gaps in the NLSY–97 cohort, we need to explore other factors. Possible examples include better measures of non-cognitive and social skills and more detailed measures of neighborhood characteristics at the census track level.[21] In the next section, I examine one of the possible factors, the school-to-work transition.

[21] I am applying to gain access to the census track geocode files of the NLSYs. With access to these data, I will be able to measure neighborhood at a more detailed level.

Role of the School-to-Work Transition

As shown in Table 3.2, compared to white men, Black men in both cohorts fell substantially behind in their very first year post-schooling. It has been documented widely in the literature that performance in 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).[22] If Black disadvantage in the school-to-work transition persists through early careers, it may help explain some of the unexplained racial labor market gaps observed in the sixth to eighth years.

Figures 3.1–3.3 reveals some suggestive patterns. In the NLSY–79 cohort, the Black-white gap in the transition (defined as the first year post-schooling) shows some convergence over the first four to five years, especially in employment outcomes. In the NLSY–97 cohort, there was much less convergence, and the initial racial gaps either largely persisted or grew over the early career years. Although this is not a causal estimate, it does suggest that the impact of the school-to-transition seems to be especially relevant for the NLSY–97 cohort and accounting for racial differences in the transition process may help explain the residual racial gaps in employment and earnings over the sixth to eighth years.

In the DFL decomposition, I estimate the explanatory power of racial differences in the school-to-work transition *conditioning on* measured racial differences in education and skills, family background, and childhood neighborhood characteristics. Table 3.5 presents the results, where I try different measures of the school-to-work transition. I first measure transition with 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, 50 weeks or more).

To measure the school-to-work transition with more exogenous variations, I also use the

[22]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.

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 UR from Local Area Unemployment Statistics (LAUS). The UR at entry state-year provides a different and presumably more exogenous measure of one's school-to-work transition status. In the NLSY–97 cohort, when the LAUS data have more detailed data at the *county* level, I also construct UR at entry county-year as another measure of transition.[23]

In the NLSY–97 cohort, racial differences in transition, as measured by weeks worked in the first year post-schooling, accounts for close to 10% of the racial labor market gaps, conditioning on measured racial differences in education and skills, family background, and childhood neighborhood characteristics. When county UR is used as the measure of transition, it accounts for 7%–8% of the racial gaps in employment and earnings, conditioning on the three pre-market factors. When first-year weeks worked *and* county UR are included together, the explanatory power of transition goes up to about 16%. Using state UR, which is a less accurate measure, does not achieve an explanatory power as close as that of county UR.

In the NLSY–79 cohort, conditioning on racial differences in the pre-market characteristics, accounting for racial differences in transition adds no extra explanatory power, no matter if I measure transition with weeks worked or state UR. In particular, as shown in the top panel of Table 3.5, the explanatory power of transition turns out to be negative in the DFL decomposition, as racial differences in the three pre-market factors have already *more than fully* accounted for racial differences in the school-to-work transition.[24]

Taking the results of the two cohorts together, it is clear that racial differences in the school-to-work transition do help explain some of the racial labor market gaps that cannot

[23]County-level UR data in LAUS go back to the year 1990 and are unavailable for most of the NLSY–79 sample years.

[24]In other words, the negative estimate means that given the measured racial differences in education and skills, family background, and childhood neighborhood characteristics, we would have predicted the racial transition gap to be even larger than it actually was in the NLSY–79 data.

be explained by the three pre-market characteristics in the NLSY–97 cohort. When interpreting the role of transition, it is important to keep in mind that the transition measures I use, weeks worked in the first year post-schooling or county (state) UR at entry time, could be subject to selection, even after controlling for education and skills, family background, and childhood neighborhood characteristics.

For example, the fact that Black men completed schooling (and entered the labor market) at a location and time with higher UR can be due to complicated reasons beyond simply bad luck. Past research has documented 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 underperformance 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 for work.

Labor Market Discrimination against Black Men

Racial discrimination has long existed in various facets of the U.S. labor market (the literature is surveyed by Altonji and Blank (1999)). For a more recent cohort close in age to the NLSY–97 cohort, Chetty et al. (2020) shows descriptively that Black men who grow up in places with greater racial bias among whites end up earning less as adults.[25]

Although I do not include a direct measure of discrimination in my decomposition analysis, there are at least three channels through which discrimination affects my results. First,

[25]The authors employ two measures of racial bias. The first is a test for implicit racial bias, available at the county level from the Race Implicit Association Database. The second is the Racial Animus Index created by Stephens-Davidowitz based on Google searches, available at the media market level (more aggregate than county).

if Black men receive lower labor market returns (such as less working time and/or lower wages) to the same pre-market characteristics due to discrimination, the racial differences in labor market returns will be left in the unexplained residuals in the decomposition. If labor market discrimination has become a more serious issue for the younger cohort than for the older cohort, it could explain, at least partially, why a greater share of the racial labor market gaps in the NLSY–97 cohort are left unexplained by racial differences in the pre-market characteristics.

Second, labor market discrimination can have a feedback effect on skill investment decisions. If it is anticipated that their skills will not be rewarded fairly in the labor market, Black men and Black families may underinvest in education and skills prior to labor market entry. As a result, some of the observed racial gaps in education and skills could be partially due to the feedback effect of labor market discrimination.

Third, as discussed in the previous section, discrimination could also contribute to racial differences in the school-to-work transition. It may have taken Black men more time to get the first job because they faced discrimination in the hiring process. And, as discussed above, part of the reason why Black men entered the labor market with a worse location and timing could be that discrimination prevented or discouraged them from searching for jobs at places with more opportunities.

3.3.2 Central Role of Education and Skills

In addition to the overall explanatory power of pre-market characteristics, my second main finding concerns the role of education and skills in explaining racial gaps in employment and earnings. Neal and Johnson (1996) show that in the NLSY–79 cohort, skills, especially cognitive skills measured by AFQT score, play the central role of explaining racial wage gaps. The second chapter of this dissertation shows that in the NLSY–97 co-

hort, education and skills remain the key explanatory factor of racial gaps in labor market outcomes, and their explanatory power largely persists even after controlling for racial differences in family background and childhood neighborhood characteristics.

The focus and contribution of my analysis here is to make a comparison across cohorts and examine how the role of education and skills (as well as family background and childhood neighborhood) has changed from the NLSY–79 cohort to the NLSY–97 cohort. In the last three columns of Table 3.3, I present the share of the racial employment and earnings gaps that can be explained by racial differences in each one of the three pre-market factors. Two patterns are worth discussion.

First, as a confirmation of past studies (Neal and Johnson, 1996; the second chapter of this dissertation), education and skills, among all measured pre-market factors, turn out to play a key role in explaining the racial gaps in employment and earnings *within* each cohort. In the NLSY–79 cohort, racial differences in education and skills *alone* explain 66%–67% of the racial employment and earnings gaps in the DFL decomposition. As a comparison, family background alone explains 51%–67% and childhood neighborhood alone explains 16%–40% of the racial labor market gaps in the DFL decomposition. In the NLSY–97 cohort, racial differences in education and skills alone explain 27%–30% of the racial labor market gaps, while family background and neighborhood characteristics each explain 20%–34% and 1%–9% of the racial gaps alone, respectively. The Oaxaca-Blinder decomposition reveals a similar pattern: education and skills show an explanatory power that is either the largest, or close to the largest, among the three observed pre-market factors.[26]

A further question is whether the central role of education and skills is driven by for-

[26]When added together in the decomposition, the *overall* explanatory power of the three pre-market factors is lower than the sum of the *sole* explanatory power of these factors when added alone. This is because education and skills, family background, and childhood neighborhood characteristics are generally positively correlated with each other.

mal schooling (measured by highest grade completed) or cognitive skills (measured by the AFQT score).[27] Table 3.6 shows the explained share of racial labor market gaps when *only* highest grade completed or AFQT score is included in the DFL decomposition. When highest grade completed is added alone, it accounts for 17%–28% of the racial gaps in the NLSY–79 cohort and 12%–14% of the racial gaps in the NLSY–97 cohort, which are less than half of the explanatory power when the full skill set is included. When AFQT score is added alone, it achieves an explanatory power that is at least close to the full skill set in both the NLSY–79 and the NLSY–97 cohorts. The results show clearly that the key role of education and skills in explaining racial labor market gaps in both cohorts mainly comes from measured racial gaps in cognitive skills.

The second key pattern from the last three columns of Table 3.3 is that racial differences in education and skills explain a greater share of the racial labor market gaps in the NLSY–79 cohort than in the NLSY–97 cohort. A similar falling explanatory power is also observed for family background and childhood neighborhood characteristics, so this pattern does not seem to be unique to education and skills.

There are at least two possible reasons why the explanatory power of education and skills has fallen across cohorts. First, as discussed above in Table 3.1, the racial gap in cognitive skills, measured by AFQT score percentile, has fallen substantially and statistically significantly from the NLSY–79 cohort to the NLSY–97 cohort. If the returns to cognitive skills were to stay 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. Recall that in Table 3.1, the racial gap has *increased* insignificantly for highest grade completed and significantly for social test score. Because the explanatory power of education and skills is primarily driven by AFQT score, rather than by highest grade completed or social test

[27] In both cohorts, racial differences in measured non-cognitive and social skills alone explain a negligible share of the racial gaps in employment and earnings; therefore I did not present the results for non-cognitive or social skills in Table 3.6.

score, the falling racial gap in AFQT score is playing the dominant role here.

Second, if labor market returns to education and skills were to decrease across cohorts, the explanatory power of education and skills would have also decreased (assuming that the 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) finds 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) shows that the former two studies rely on a strong assumption of constant skill prices and the assumption does not 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.[28]

Exploration of the Relationship between Skills and Family and Neighborhood Factors

It is important to understand the relationship between education and skills and racial differences in family background and childhood neighborhood characteristics. The variables of education and skills in the NLSY datasets are measured over childhood and/or early adulthood, and the observed racial differences in education and skills could have already been exposed to family and neighborhood influences.[29]

[28] Although there is a lack of strong supportive evidence for changing returns to education and skills across cohorts, it is worth noting that I also cannot rule out the possibility that the importance of other factors, such as family background, childhood neighborhood, or labor market discrimination, has changed across cohorts. This topic merits a formal exploration in future research.

[29] 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

To examine how much of the explanatory power of education and skills comes from racial differences in family background and childhood neighborhood, I exploit the sequential nature of the DFL decomposition. Intuitively, under the DFL decomposition, each factor can be added in a sequential manner, and the explanatory power of factors added later is estimated conditioning on racial differences in the factors added earlier. For example, if I add family background and childhood neighborhood in the sequence before education and skills, then the explanatory power of education and skills will be estimated after conditioning on racial differences in family and neighborhood characteristics.

In Table 3.7, I explore three different sequential orderings, where I add education and skills as the first factor, the second factor, and the last factor. Meanwhile, I always add family background before childhood neighborhood in the sequence, under the assumption that family background is determined before (so more exogenous than) childhood neighborhood in the skill accumulation process.[30]

In the NLSY–79 cohort, the explanatory power of education and skills falls from 66%-67% to 24%–39% when conditioning on racial differences in family background, and it further falls to close to zero when conditioning on racial differences in both family background and childhood neighborhood characteristics.[31] Compared to the NLSY–79 cohort, the explanatory power of education and skills in the NLSY–97 cohort falls by a much smaller degree when conditioning on family and neighborhood characteristics. When added alone,

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. 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 not directly comparable between the two NLSY cohorts. The findings in this chapter are quantitatively robust when excluding non-cognitive and social scores from the analysis.

[30] As discussed earlier, the neighborhood-level characteristics that I include capture *both* the average characteristics of individuals and families living in the neighborhood and the effect of neighborhood per se. Adding family variables before neighborhood variables in the sequence will help isolate the true effect of neighborhood.

[31] The overall explanatory power of all three pre-market factors does not change with the specific sequential ordering.

education and skills account for 27%–30% of the racial employment and earnings gaps in the NLSY–97 cohort. After conditioning on racial differences in family background, the explanatory power of education and skills falls slightly to 21%–25%. After further conditioning on racial differences in childhood neighborhood, the explanatory power of education and skills again falls only slightly to 17%–25%.

The different results between the two cohorts mean that the relationship between education and skills with family and neighborhood characteristics has changed substantially. In the older cohort, the documented central role of education and skills in shaping racial labor market gaps comes almost completely from racial differences in family background and childhood neighborhood characteristics. From a policy perspective, this suggests that a promising pathway to reduce racial gaps in labor market outcomes is family- and neighborhood-based programs that aim to reduce Black disadvantages in family and neighborhood characteristics as measured in the NLSY datasets, such as family income, family structure, and socioeconomic conditions of childhood county of residence.

On the other hand, in the younger cohort much of the explanatory power of education and skills persists even after accounting for measured racial differences in family and neighborhood characteristics. Note that family and neighborhood characteristics together account for about 30% of the racial employment and earnings gaps in the NLSY–97 cohort, as shown in the bottom two rows of Table 3.7.

3.3.3 Decomposition of Racial Gaps across the Distribution

My analysis until now has focused on racial employment and earnings gaps observed at the mean. In this section, I briefly present the DFL decomposition results of racial gaps observed at different parts of the employment and earnings distributions.[32] Specifically, I

[32]The DFL method estimates the whole counterfactual distributions of employment or earnings for white men if they had the same pre-market characteristics as Black men. In principle, the method can be used to

focus on racial gaps measured at the 25th percentile, the median, and the 75th percentile of the employment and earnings distributions.

Table 3.8 replicates the structure of Table 3.3 but with racial gaps at different parts of the distribution instead of the mean. The employment distribution (average weeks worked per year in the sixth to eighth years) is highly right-skewed, and more than 25% of both Black and white men worked for 52 weeks, on average, over the sixth to eighth years in both the NLSY–79 and the NLSY–97 cohorts. As a result, the racial employment gap is zero at the 75th percentile, and I do not include it in Table 3.8.

The basic patterns observed at the mean hold qualitatively at different parts of the distributions. First, the overall explanatory power of education and skills, family background, and childhood neighborhood together is higher in the NLSY–79 cohort than in the NLSY–97 cohort. In particular, racial differences in the three pre-market factors more than fully account for racial gaps at the 25th percentile and the median of the employment distribution in the NLSY–79 cohort. The only exception is that racial differences in pre-market factors together account for 48% of the racial gap at the 75th percentile of the earnings distribution in the NLSY–97 cohort, which is greater than the overall explanatory power of pre-market factors in the NLSY–79 cohort (38%).

Second, education and skills still have the highest, or close to highest, explanatory power among all pre-market characteristics. In particular, racial differences in education and skills *alone* more than fully explain the racial employment gaps at the 25th percentile and the median in the NLSY–79 cohort. Third, the sole explanatory power of education and skills, as well as that of family or neighborhood characteristics, has fallen from the NLSY–79 cohort to the NLSY–97 cohort. The only exception is, again, at the 75th percentile of the earnings distribution, where education and skills alone explain 54% of the gap in the NLSY–97 cohort and 38% of the gap in the NLSY–79 cohort, and family background

decompose racial gaps at any part of the employment or earnings distribution.

characteristics alone explain 33% of the gap in the NLSY–97 cohort and 25% of the gap in the NLSY–79 cohort.

3.4 Conclusion

How have the racial labor market gaps among young men changed across cohorts, and how have the underlying drivers of racial gaps changed? In this chapter, I answer these questions with the help of two similarly constructed and nationally representative samples of young Americans, the NLSY–79 and the NLSY–97 cohorts. I find that racial gaps in employment and earnings observed at specific time points in the first eight years post-schooling are not statistically distinguishable between the two cohorts. However, the shapes of employment and earnings trajectories have changed dramatically across cohorts for both races.

Using a semi-parametric decomposition, I further show that compared to the NLSY–79 cohort, measured racial differences in pre-market education and skills, family background, and childhood neighborhood in the NLSY–97 cohort explain a lower share of the racial labor market gaps observed over the sixth to eighth years post-schooling. This finding suggests that racial differences in unobservable factors have played a more important role in shaping racial labor market gaps in the younger cohort.

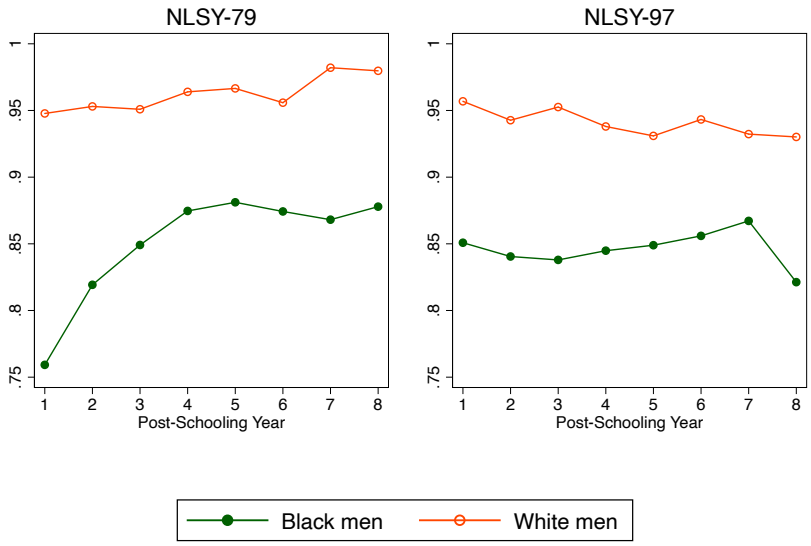
I provide some speculative evidence on what the unobservable factors can be. I show that Black disadvantage in the school-to-work transition, measured with weeks worked in the first year post-schooling or county UR in the year of labor market entry, explains a quantitatively important share of the racial labor market gaps in the NLSY–97 cohort. In contrast, I find that more detailed measures of childhood neighborhood characteristics add little extra explanatory power to the existing pre-market factors. I argue that racial discrimination in the labor market can be an important source of the unobservable factors.

More research needs to be done to understand how labor market discrimination against Black men has evolved across cohorts and the potential consequences on racial gaps in labor market outcomes.

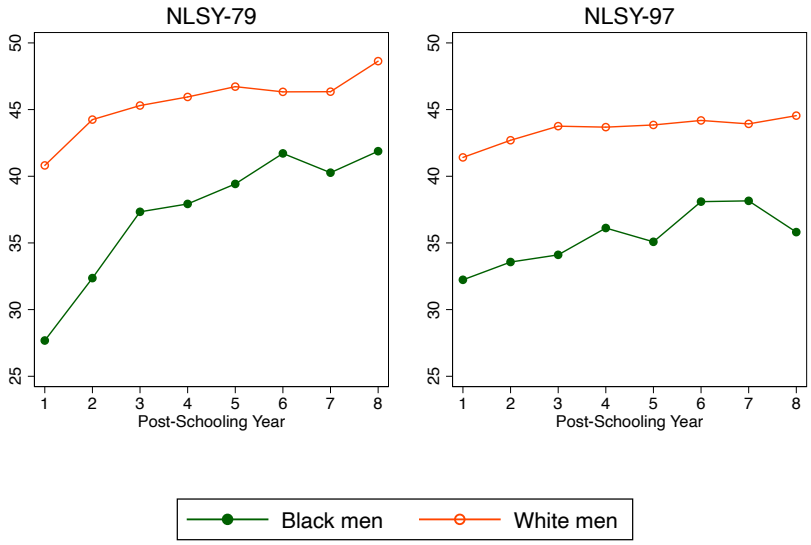
Among observable pre-market factors, I show that education and skills appear to have played the central role in explaining racial labor market gaps in both cohorts, which is consistent with the findings of Neal and Johnson (1996) and the second chapter of this dissertation. The explanatory power of education and skills in the NLSY–79 cohort comes almost fully from racial differences in measured family and neighborhood characteristics, while much of the explanatory power of education and skills in the NLSY–97 cohort cannot be attributed to family and neighborhood characteristics as they are measured in the data. This suggests that in addition to family- and neighborhood-based programs, polices that aim at reducing racial labor market gaps need to pay more attention to Black disadvantage in the accumulation of education and skills. In future work, I plan to further investigate what has contributed to the observed racial gap in education and skills in the younger cohort and what it implies for designing more effective public policies to reduce racial labor market gaps for future cohorts to come.

Figure 3.1: Career Trajectories: Any Employment and Weeks Worked

(a) Any Employment



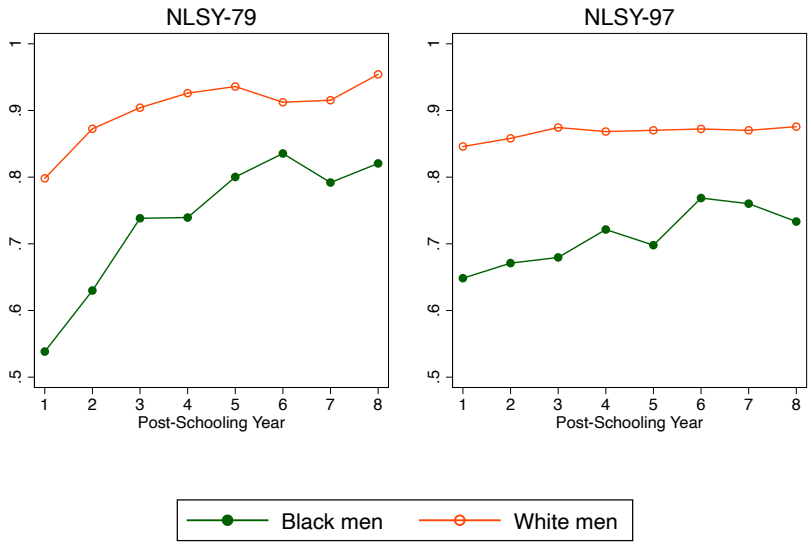
(b) Weeks Worked



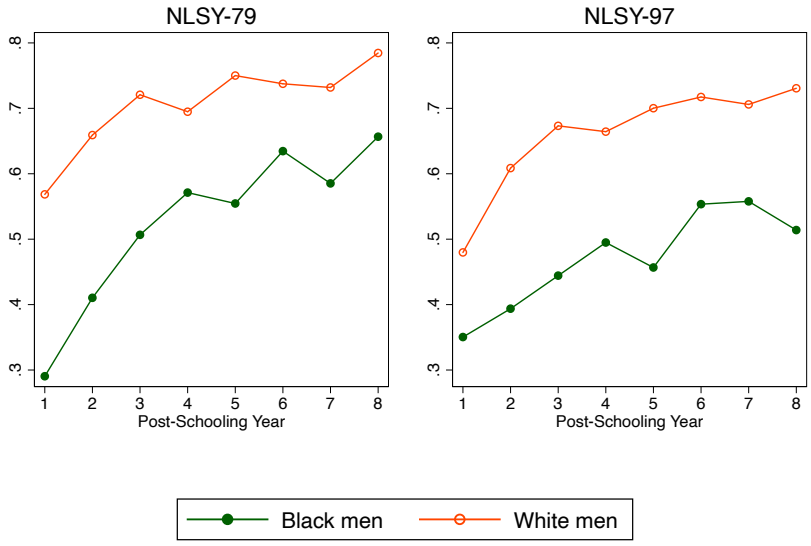
Notes: Both NLSY–79 and NLSY–97 samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used.

Figure 3.2: Career Trajectories: Worked Half Year and Full Year

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



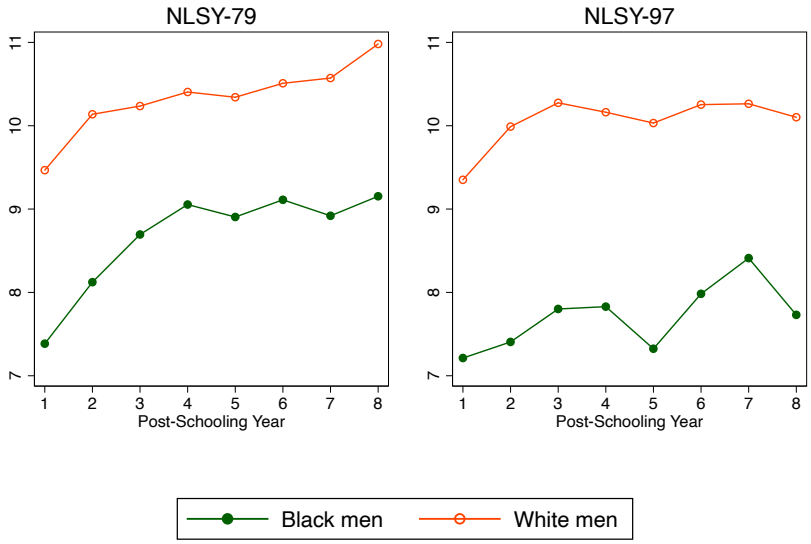
(b) Worked Full Year ($\geq$ 50 weeks)



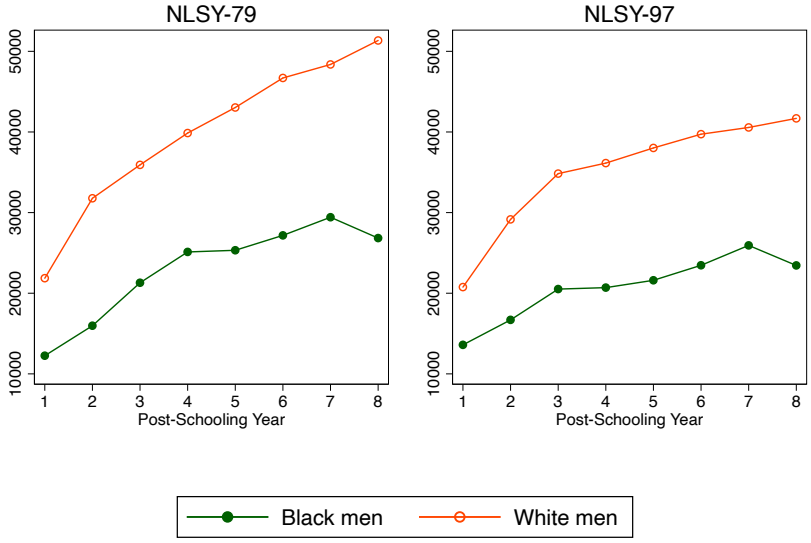
Notes: Both NLSY–79 and NLSY–97 samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used.

Figure 3.3: Career Trajectories: Annual Earnings

(a) Log Annual Earnings



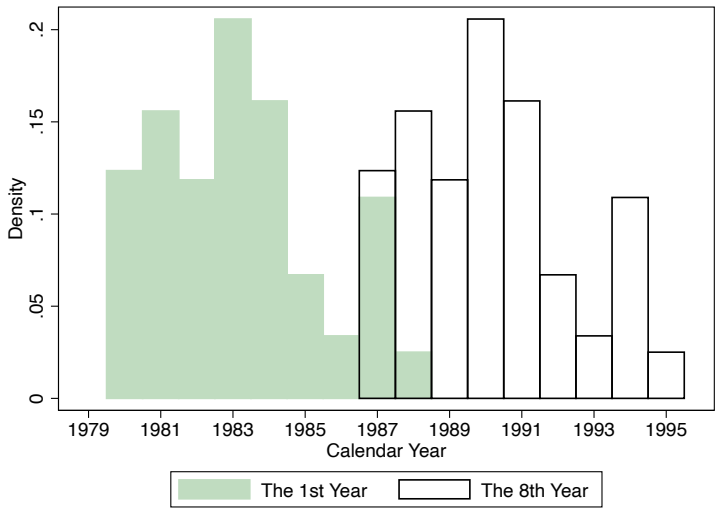
(b) Annual Earnings



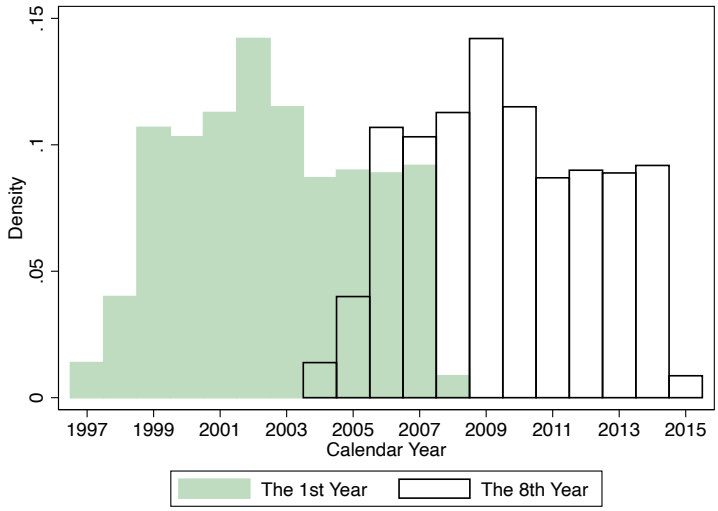
Notes: Annual earnings are adjusted to 2013 dollars. The top panel takes the inverse hyperbolic sine of annual earnings. Both NLSY–79 and NLSY–97 samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used.

Figure 3.4: 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.

Table 3.1: 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 3.2: Early Career Outcomes 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)
Transition Stage (1st Year)							
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
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] NE stands for non-employment. 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 3.3: Explanatory Power of Observed Pre-Market Factors

	Racial Gap	Share Explained by Pre-Market Factors	Unexplained Residuals	Share Explained by Each Factor When Added Alone		
				Skills	Family	NBHD
	(1)	(2)	(3)	(4)	(5)	(6)
NLSY–79						
DFL						
Weeks worked per year	5.78	86%	14%	66%	51%	40%
Log avg annual earnings	1.37	77%	23%	67%	67%	16%
Oaxaca-Binder						
Weeks worked per year	5.78	95%	5%	78%	31%	31%
Log avg annual earnings	1.37	93%	7%	77%	33%	35%
NLSY–97						
DFL						
Weeks worked per year	6.90	48%	52%	27%	34%	1%
Log avg annual earnings	1.54	55%	45%	30%	20%	9%
Oaxaca-Binder						
Weeks worked per year	6.90	55%	45%	28%	31%	11%
Log avg annual earnings	1.54	49%	51%	30%	23%	13%

[1] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and NBHD stands for neighborhood. The top panel uses the NLSY–79 sample, including 444 white men and 271 Black men. The bottom panel uses the NLSY–97 sample, including 825 white men and 396 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 3.4: Explanatory Power of Additional Pre-Market Factors in NLSY–97

	Racial Gap	Share Explained by Pre-Market Factors	Unexplained Residuals	Share Explained by Each Factor When Added Alone		
				Skills	Family	NBHD
	(1)	(2)	(3)	(4)	(5)	(6)
NLSY–79						
DFL						
Weeks worked per year	5.78	86%	14%	66%	51%	40%
Log avg annual earnings	1.37	77%	23%	67%	67%	16%
Oaxaca-Binder						
Weeks worked per year	5.78	95%	5%	78%	31%	31%
Log avg annual earnings	1.37	93%	7%	77%	33%	35%
NLSY–97						
DFL						
Weeks worked per year	6.90	48%	52%	27%	34%	1%
Log avg annual earnings	1.54	55%	45%	30%	20%	9%
Oaxaca-Binder						
Weeks worked per year	6.90	55%	45%	28%	31%	11%
Log avg annual earnings	1.54	49%	51%	30%	23%	13%
NLSY–97 (with more variables)						
DFL						
Weeks worked per year	6.90	49%	51%	27%	35%	11%
Log avg annual earnings	1.54	49%	51%	30%	21%	15%
Oaxaca-Binder						
Weeks worked per year	6.90	62%	38%	28%	34%	24%
Log avg annual earnings	1.54	57%	43%	30%	26%	18%

[1] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and NBHD stands for neighborhood. The top two panels replicate Table 3.3 with the NLSY–79 sample and the NLSY–97 sample. The bottom panel uses the same NLSY–97 sample but adds family and neighborhood variables that are available only in the NLSY–97 cohort, including whether the respondent's mother is a teenage mom, homeownership status, and neighborhood quality measures created by Chetty and Hendren (2018b). 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 3.5: Role of School-to-Work Transition Conditional on Pre-Market Factors

	Racial Gap	Share Explained by Racial Differences in				
	(1)	(2)	(3)	(4)	(5)	(6)
NLSY–79						
		Weeks worked in 1st Year	State UR at Entry	Weeks & State UR		
Weeks worked per year	5.78	–10%	–26%	–26%		
Log avg annual earnings	1.37	–19%	–12%	–26%		
NLSY–97						
		Weeks worked in 1st Year	State UR at Entry	Weeks & State UR	County UR at Entry	Weeks & County UR
Weeks worked per year	6.90	9%	2%	10%	7%	16%
Log avg annual earnings	1.54	8%	1%	9%	8%	16%

[1] DFL stands for the DiNardo-Fortin-Lemieux decomposition, and UR stands for unemployment rate. The top panel uses the NLSY–79 sample and the bottom panel uses the NLSY–97 sample. Sample weights are used.

[2] Column 2 measures the school-to-work transition with a flexible vector of weeks worked in the first year post-schooling. Column 3 instead uses UR in one's state of residence at the labor market entry year. Column 4 includes both of these two measures. In the bottom panel (NLSY–97), I add two more columns that replace state UR with county UR. County-level UR data, provided by the Local Area Unemployment Statistics (LAUS) program, go back to 1990 and is unavailable for most of my NLSY–79 sample years. In all columns, transition measures are added after family background, childhood neighborhood, and education and skills in the DFL decomposition.

[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.

Table 3.6: Explanatory Power of Formal Schooling versus Measured Cognitive Skills

	Racial Gap	Share Explained by Racial Differences in		
	(1)	(2)	(3)	(4)
NLSY–79				
		Full Skill Set	Highest Grade Completed	AFQT score
Weeks worked per year	5.78	66%	28%	107%
Log avg annual earnings	1.37	67%	17%	83%
NLSY–97				
		Full Skill Set	Highest Grade Completed	AFQT score
Weeks worked per year	6.90	27%	14%	28%
Log avg annual earnings	1.54	30%	12%	30%

[1] AFQT stands for the Armed Forces Qualification Test. The top panel uses the NLSY–79 sample, including 444 white men and 271 Black men. The bottom panel uses the NLSY–97 sample, including 825 white men and 396 Black men. Both samples are balanced panels of men who have completed formal schooling for at least eight years. Sample weights are used. Column 2 replicates Table 3.3 and uses the full set of skill measures, which includes highest grade completed, AFQT score, and non-cognitive and social test scores. Column 3 includes only the highest grade completed in the skill set. Column 4 includes only AFQT score in the skill set. In all three columns, *only* skill measures are included in the DFL decomposition, while family background, childhood neighborhood, and transition variables are not included.

[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 inverse hyperbolic sine.

Table 3.7: Sequential DFL Decomposition Results

	Racial Gap	Share Explained by Each Factor in Sequential Ordering			Unexplained Residuals
	(1)	(2)	(3)	(4)	(5)
NLSY–79					
Sequential Ordering I		Skills	Family	NBHD	
Weeks worked per year	5.78	66%	25%	–5%	14%
Log avg annual earnings	1.37	67%	24%	–14%	23%
Sequential Ordering II		Family	Skills	NBHD	
Weeks worked per year	5.78	51%	39%	–5%	14%
Log avg annual earnings	1.37	67%	24%	–14%	23%
Sequential Ordering III		Family	NBHD	Skills	
Weeks worked per year	5.78	51%	27%	8%	14%
Log avg annual earnings	1.37	67%	12%	–2%	23%
NLSY–97					
Sequential Ordering I		Skills	Family	NBHD	
Weeks worked per year	6.90	27%	28%	–7%	52%
Log avg annual earnings	1.54	30%	14%	10%	45%
Sequential Ordering II		Family	Skills	NBHD	
Weeks worked per year	6.90	34%	21%	–7%	52%
Log avg annual earnings	1.54	20%	25%	10%	45%
Sequential Ordering III		Family	NBHD	Skills	
Weeks worked per year	6.90	34%	–3%	17%	52%
Log avg annual earnings	1.54	20%	11%	25%	45%

[1] NBHD stands for neighborhood, and DFL stands for the DiNardo-Fortin-Lemieux decomposition. The top panel uses the NLSY–79 sample and the bottom panel uses the NLSY–97 sample. 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 3.8: Explanatory Power of Pre-Market Factors across the Distribution

	Racial Gap	Share Explained by Pre-Market Factors	Unexplained Residuals	Share Explained by Each Factor When Added Alone		
				Skills	Family	NBHD
	(1)	(2)	(3)	(4)	(5)	(6)
NLSY–79						
Employment Distributions						
25th percentile	8.33	164%	–64%	112%	60%	40%
Median	1.33	300%	–200%	150%	50%	25%
Earnings Distributions						
25th percentile	0.65	34%	66%	75%	57%	4%
Median	0.70	45%	55%	58%	27%	–1%
75th percentile	0.56	38%	62%	38%	25%	–8%
NLSY–97						
Employment Distributions						
25th percentile	15.00	31%	69%	27%	33%	0%
Median	6.00	11%	89%	8%	11%	0%
Earnings Distributions						
25th percentile	1.45	23%	77%	15%	17%	3%
Median	0.68	24%	76%	20%	20%	–3%
75th percentile	0.41	48%	52%	54%	33%	–7%

[1] NBHD stands for neighborhood. The top panel uses the NLSY–79 sample, including 444 white men and 271 Black men. The bottom panel uses the NLSY–97 sample, including 825 white men and 396 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.

Chapter 4: Have the Returns to Cognitive Skills Really Declined in the U.S.?

4.1 Introduction

Understanding how human capital differences across workers affect their labor market outcomes is an essential goal of empirical research in labor economics. The early focus on modeling and estimating the return to education (e.g. Becker, 1964; Mincer, 1974) gave rise to consideration of how underlying dimensions of skill, such as cognitive ability, affect the accumulation and estimation of returns to education (e.g. Griliches, 1977). In more recent decades, the focus has partly shifted to understanding and estimating the returns to these dimensions of skill themselves (e.g. Neal and Johnson 1996; Heckman, Stixrud, and Urzúa, 2006; Urzúa, 2008), rather than just treating them as a nuisance factor in estimating the returns to education. And amidst all of this, researchers have turned to understanding how dramatic changes in the structure of the U.S. labor market over the past few decades has change the returns to schooling and underlying skills (e.g. Autor, Dorn, and Hanson, 2013; Beaudry, Green, and Sand, 2016; Autor, 2019).

One of the large challenges to estimating the labor market return to any specific type of skill is that researchers only ever observe proxies for underlying skills of individuals, and these proxies may suffer from substantial measurement error (Griliches, 1986). Establishing how the return to a specific skill may have evolved over time is even more challenging, as measurement issues are compounded by the fact that the observed proxies for those skills

themselves often change over time.

Two influential studies, both using data from the two cohorts of the National Longitudinal Survey of Youth (NLSY), provide empirical evidence suggesting that the wage returns to cognitive skills have declined for young workers in the U.S. over the past 40 years (Castex and Dechter, 2014; Deming, 2017). In both of these studies, the authors use information from the Armed Forces Qualifying Test (AFQT) to measure cognitive skills across the two NLSY cohorts. Casual observation might suggest that because AFQT test scores are collected and utilized for both NLSY cohorts, measurement issues with using AFQT as a proxy for cognitive ability should be the same across cohorts, alleviating concerns that might have arisen if different cognitive ability test scores had been used. Castex and Dechter (2014) argue that returns to education have increased while returns to cognitive ability have decreased. Deming (2017) concludes that returns to cognitive skills have declined, but that returns to social skills have increased. Both papers use the empirical conclusions that they make from the NLSY data to try to better understand *why* these returns have changed, and although the two papers have different emphases, both end up concluding that technological change has fundamentally altered the relationship between skills and labor market outcomes such that cognitive ability is less productive, and thus the returns to cognitive skill has fallen.

In this paper, we revisit the question of whether the returns to cognitive ability have changed. First, we point out that if technology is indeed driving the changing wage return to cognitive ability, it should also be driving investments in cognitive ability, an implication that the previous papers have not considered formally. These changing investments in cognitive ability should cause a shift in the population distribution of cognitive ability over time. We examine the theoretical implications of this in the first section of the paper by specifying and simulating a simple model of skill accumulation with changing technology, demonstrating that we can rationalize both a change in the wage return to cognitive ability

and a change in the the underlying distribution of cognitive ability. We then introduce a Yitzhaki (1996) decomposition to demonstrate the relationship between the distribution of cognitive ability and the estimation of its wage return in the labor market in the typical setting of a linear regression where the log of wages is regressed on a measure of cognitive ability.

In the second section of the paper, we document (as in Altonji, Bharadwaj, and Lange, 2012) the change in distribution over time in what we call the AAFQT, a measure of cognitive ability derived from the Armed Forces Qualifying Test (AFQT) that has been collected and concorded in the two cohorts of the National Longitudinal Survey of Youth (NLSY). Empirically, we show that the changing distribution of AAFQT over the cohorts, while somewhat more complex than our simple model suggests, is consistent with a model of changing investments in cognitive skill in the wake of a fall in the productivity of cognitive ability.

We then perform the Yitzhaki decomposition of the ordinary least squares (OLS) estimates of the returns to cognitive skill separately for each NLSY cohort, using the typical log linear functional form relationship between wages and cognitive skill. We recover the OLS estimates of the wage return for each cohort, but we also demonstrate how the OLS estimates are affected by the distribution of AAFQT via an analysis of the weights from Yitzhaki's formulation. We show that the changing distribution of AAFQT over time plays a key role in the empirical finding of declining OLS returns to AAFQT. Given this finding, we end with a discussion of whether the observed changes in the test score measure may be driven by measurement error, rather than by true changes in the distribution of cognitive ability.

4.2 The Conceptual Framework

Consider a general process generating a labor market outcome, Y, of the form:

$$Y = \varphi(A, H) \tag{4.1}$$

where A denotes technology, H represents human capital. For sake of clarity, we omit other characteristics. A long-standing literature in economics has confirmed that:

$$\frac{\partial Y}{\partial H} = \frac{\partial \varphi(A, H)}{\partial H} \geq 0. \tag{4.2}$$

This basic setting has, for decades, formed the basis of much of what labor economists study. While an individual's human capital, H, was initially specified as a univariate measure–often years of schooling (e.g. Becker, 1964; Mincer, 1974)–more recently it has been thought of as a multi-dimensional measure of skills, consisting (depending of data availability) of years of schooling and a vector of skills (often cognitive skills and sometimes non-cognitive skills and social skills).

Given the implications of technological change for the labor markets (Acemoglu, 2002), in this paper we examine whether the relationship between skills and labor market outcomes has also changed as a result of this process. In particular, the parameter of interest is:

$$\frac{\partial^2 Y}{\partial A \partial H} = \frac{\partial^2 \varphi(A, H)}{\partial A \partial H}. \tag{4.3}$$

Without data on A, researchers have turned to time variation as a source of identification for equation (4.3). Castex and Dechter (2014) and Deming (2017) estimate versions of this expression using this approach. As the relevant labor market outcome they use hourly wages where the data are drawn from the 1979 and 1997 cohorts of the National Longi-

tudinal Survey of Youth (NLSY). Both papers provide empirical evidence suggesting that the wage returns to dimensions of human capital have changed across the cohorts. On conceptual grounds, however, it is necessary to consider the potential endogeneity of skills before interpreting the sign of expression (4.3) as evidence of this claim. In addition, just the presence of (potentially non-classical) measurement error in proxies for skill can lead to incorrect conclusions.

Consider first the role of skill formation and assume a structural association between A and H, $H = H(A)$. In a dynamic setting, this could be the result of individual's human capital investments responding to technology. To simply the analysis, we assume H is univariate. Thus, the total impact of technology is:

$$\frac{dY}{dA} = \frac{\partial Y}{\partial H} \times \frac{dH}{dA} + \frac{\partial Y}{\partial A},$$

from which we obtain:

$$\frac{d^2Y}{dA^2} = \left[\frac{\partial^2 Y}{\partial H^2}\frac{dH}{dA} + \frac{\partial^2 Y}{\partial A \partial H}\right]\frac{dH}{dA} + \frac{\partial Y}{\partial H}\frac{d^2H}{dA^2} + \frac{\partial^2 Y}{\partial A \partial H}\frac{dH}{dA} + \frac{\partial^2 Y}{\partial A^2}.$$

If we further impose that $\frac{\partial^2 Y}{\partial H^2} = \frac{\partial^2 Y}{\partial A^2} = 0$ (linearity), the sign of (4.3) is determined by the sign of the expression:

$$\frac{\frac{d^2Y}{dA^2} - \frac{\partial Y}{\partial H}\frac{d^2H}{dA^2}}{dH/dA} \tag{4.4}$$

which illustrates how technology and human capital accumulation determine the way skill premium evolves with technological change.[1]

[1] On top of the analysis here, selection into occupation amplifies the difficulties of empirically identifying the sign of (4.3). To see this, consider a two-sector Roy model. Let A_1 and A_2 be the respective technologies in each sector. Individuals self-select into 1 or 2. Let D_2 be a dummy indicating sector 2 has been selected. Thus, the outcome becomes:

$$Y = \varphi(A_1, H) + D_2\left[\varphi(A_2, H) - \varphi(A_1, H)\right], \tag{4.5}$$

4.2.1 A simple illustrative model

To better illustrate the economic forces driving the relationship between technology and human capital investments, we consider an adaptation of the "q" theory of investment (Tobin and Brainard, 1976) to our setting. We emphasize at the outset that our goal here is not to fully fit the model to the data we use, but rather to point out in a simple framework that if cognitive ability is the result of an investment process, technological change will affect the population distribution of cognitive ability and its return in the labor market. Let I_t denote human capital investments, H_t the stock of human capital, and $I_t = (dH_t/dt)/H_t$. Let $q(I)$ denote the cost of adding human capital ($q'(I) > 0, q''(I) > 0, q(0) = 0, q'(0) = 1$), ρ the discount factor, r the marginal productivity of human capital (in units of output), and w the wage (per unit of output). The agent solves:

$$Max_{I_t} \int_0^\infty w(r - q(I_t))H_t e^{-\rho t} dt.$$

The optimal I^* must satisfy $\frac{r-q(I^*)}{\rho - I^*} = q'(I^*)$ (Uzawa, 1969). Assuming a quadratic cost function and a discount factor of 0.04, a 10% decrease in the marginal productivity of human capital r (from $r = 1.1$ in period 1 to $r = 1$ in period 2) would drive down optimal investments I^* from 1.8% to 1.6%. This could be the result of technological change generating the mapping from A to H described above.

Assuming in the initial condition that H_0 follows a beta distribution $B(5, 1.5)$, Figure 4.1 depicts the distribution of the stock human capital under two regimes (period 1 and period 2) with ten years of endogenous investments. After the change in productivity, the distribution of human capital shifts leftward in period 2.

and the sign of (4.3) now depends on how ability and technology alter the occupational decision-making and potential outcome gains. Accounting for selection demands becomes essential. Despite the fact that Deming (2017) acknowledges the issue, the literature has overlooked its implications.

4.2.2 The Yitzhaki Decomposition

Despite these complexities, the evidence on the sign of expression (4.3) comes from Least Squares results obtained from linear regression models of the form:

$$Y = E[Y|H,A] + \varepsilon = \alpha(A) + \beta(A)H + \varepsilon. \tag{4.6}$$

In order to estimate the equation, one must choose Y, and one must find measures of A and H. Deming (2017) and Castex and Dechter (2014) estimate versions of this equation where they specify log wages as Y, they use time as a proxy for A, and they use a measure of cognitive skill that is derived from Altonji et al. (2012). They both report a decreasing β_A between NLSY–79 and NLSY–97, both concluding the returns to (cognitive) ability have decreased over time.

In light of expression (4.4), we explore the potential channels behind this conclusion, by implementing a Yitzhaki decomposition (Yitzhaki, 1996). Let $B_A(h) = E(Y|A, H = h)$ be the regression curve and $b_A(h)$ be its slope, i.e. the unit treatment effect evaluated at h given A. The Ordinary Least Squares (OLS) estimate of the linear relationship between Y and H can be expressed as:

$$\beta_A^{OLS} = \int_h w_A(h)\, b_A(h)\, dh. \tag{4.7}$$

Therefore, β_{OLS} can be decomposed into unit treatment effects $b_A(h)$ and how they are weighted by $w_A(h)$ across A.[2] The weights can be expressed as:

$$w_A(h) = \frac{F_{A,H}(h)}{\sigma^2_{A,H}} \Big(E_A(H) - E_A(H|H \leq h) \Big)$$

[2]Theil (1950) and Sen (1968) exploit this formulation.

where $F_{A,H}(h)$ is the cumulative density function of H evaluated at h, $\sigma^2_{A,H}$ is the associated variance, and $E_A(\cdot)$ denotes the expectation given A. Simple algebra yields:

$$w_A(h) \quad = \quad \left[\frac{F_{A,H}(h)\Big(1-F_{A,H}(h)\Big)}{\sigma^2_{A,H}}\right] \Big(E_A(H|H>h)-E_A(H|H\leq h)\Big) \qquad (4.8)$$

$w_A(h)$ are non-negative and solely depend on the distribution of H. $F_{A,H}(h)(1-F_{A,H}(h))$ reaches its peak when $F_{A,H}(h)=0.5$, i.e. at the center of the distribution of H, and therefore so does the term in the first set of parentheses in (4.8).

Looking at the second expression in parentheses in (4.8), one sees that larger differences in the conditional expectations on either side of a given h contribute more to the OLS weights. This dispersion is essential for understanding how OLS operates. In particular, what it means is that left(right)-skewed distributions will tend to put higher (lower) weight on h's toward the bottom of the distribution. Note, however, that if the unit treatment effects $b_A(h)$ are constant, the weighting scheme of the Yitzhaki decomposition doesn't matter in practice. This will turn out to be important to our analysis later on.

We can implement the Yitzhaki decomposition in the context of our example. Given the evolution of the stock of human capital and assuming $W = w \times e^{r\times H}$, we can reproduce what a linear regression of log wages on human capital ($\ln W = \alpha + \beta H$) would yield in each period. Figures 4.2 and 4.3 present the Yitzhaki slopes and weights, respectively. We see how the increase in r affects both the weights and slopes at the same time. In this specific simulation case (and in particular, given the specific wage function), we see the slopes are constant in both periods and are lower in period 2 across the H distribution, and the weights shift leftward. The changes in slopes and weights together make the OLS estimator of the effect of human capital on wages fall by 9.1% from period 1 to period 2.

This example is quite simple, of course, and changing the model in various ways–for

example by changing the structure of the cost function to be something more complex than a simple quadratic function–will yield different investment decisions that will yield different distributions of cognitive ability, different Yitzhaki weights and unit slopes, and thus different wage returns.

Another issue of a more practical nature is measurement error. As true ability or skill might be difficult to observe, one could conceptualize any ability measure, H^*, as $H^* = H + u$. In this context, β_A^{OLS} would be biased and $w_A(h)$ and $b_A(h)$ would incorrectly characterize the true weights and unit treatment effects, respectively. Nonetheless, under classical measurement error one could still identify whether the returns to skills are increasing or decreasing across different values of A. For example, and anticipating our discussion, as casual observation might suggest that the because test scores, H^*, are collected and utilized for both NLSY cohorts, any measurement issues with using H^* as a proxy for cognitive ability, if they exist at all, should be the same across cohorts. But as we document below, a serious examination of how measured test scores have changed across the NLSY cohorts, and how those changes have affected estimated wage returns to cognitive ability, suggests that casual observation is likely wrong. As a result, one needs to be cautious in drawing firm conclusions about whether or not the return to cognitive ability has really changed. And therefore, one needs to be cautious in developing (and testing) theories for the economic processes that can drive a decline in the return. We discuss the data, including the imperfect proxies for cognitive skills, next.

4.3 NLSY–79, NLSY–97, and their imperfect proxies for cognitive skills

The National Longitudinal Survey of Youth (NLSY) has been hugely important in shaping our understanding of the U.S. labor market in the past 40 years. In particular, a tremendous amount of work has been done using NLSY–79, a nationally representative sample of

the cohort of American youths aged 14 to 22 when first surveyed in 1979. Compared to NLSY–79, less research has been conducted on NLSY–97, a younger cohort of American youth aged 12-16 when first surveyed in 1997. Now that the NLSY–97 cohort are all firmly in their thirties, it has become increasingly important to invest further in understanding the various dimensions of the labor market experiences of this cohort. Drawing comparisons between the labor market experiences of two NLSY cohorts, when conducted appropriately, will also help us understand how the U.S. labor market has evolved over time.

Given the recent availability of education, migration and especially labor market data of the NLSY–97 cohort, there has been an emerging series of papers that compare the two NLSY cohorts along different dimensions (Dillion and Smith 2020; Johnson and Schulhofer-Wohl 2019; Altonji, Bharadwaj, and Lange, 2012; Castex and Dechter, 2014; Deming, 2017). Among them, Altonji, Bharadwaj and Lange (2012), Castex and Dechter (2014), and Deming (2017) focus on the labor market outcomes of the two cohorts of young Americans and their findings together have shaped the narrative of changing skills and changing returns to skills in the U.S. labor market. Given the importance and influence of the three papers, we reconsider their findings in this paper.

As one of the earliest papers that conduct cross-cohort comparisons of the NLSY–79 and the NLSY–97, Altonji, Bharadwaj and Lange (2012) compare the distribution of skills between the two cohorts and study its implication for wage inequality. They carefully document the challenges associated with making cross-cohort comparisons using NLSY and construct the samples and variables in a way that best allows for concordance between the two cohorts. Based on a number of assumptions, they conclude that the skill distribution has widened from NLSY–79 to NLSY–97. In particular, they find that the distribution of AFQT score, as a measure for cognitive skills, has widened from NLSY–79 to NLSY–97. Since Altonji, Bharadwaj and Lange (2012), there have been two influential papers that study how the wage returns to skills have changed across the two NLSY cohorts of young

adults in their early careers. Castex and Dechter (2014) focus on estimating changes in the skill price of AFQT, while Deming (2017) focuses on changes in the price of measures of social skills but he also estimates the changes in the skill price of AFQT. Both papers estimate conventional linear hedonic wage functions with constant skill prices. Despite the differences in choices of samples and model specifications, both papers find that the OLS estimate of wage returns to AFQT has declined from NLSY–79 to NLSY–97. [3]

The AFQT score has been widely used by researchers as a measure of cognitive skills (see, e.g. Neal and Johnson 1996; Heckman, Stixrud, and Urzúa, 2006; Urzúa, 2008). The score is constructed based on four sections of the Armed Services Vocational Aptitude Battery (ASVAB), a test adopted by the U.S. military for determining various aspects of enlistment (Defense Manpower Data Center, 2006).[4] Individuals selected to participate in both the 1979 and 1997 NLSY cohorts also took the ASVAB. Indeed, the administration of the ASVAB was facilitated by the U.S. military in order to create an estimate of the youth population test distribution so that the scores of military recruits could be normed against the population. However, there were two fundamental differences in the test format and in the administration of the test across the two cohorts. First, while the NLSY–79 respondents took a paper-based test, the NLSY–97 respondents took a computer-based tests that incorporates Item Response Theory (IRT), so that not all respondents answer all questions. Second, the NLSY–79 respondents were ages 15-23 when they took the test, while the NLSY–97 respondents were 12-18.

Altonji et al. (2012) were the first to conduct an in-depth cross-cohort analysis of the

[3]Unlike Castex and Dechter (2014) and Deming (2017), Weinberger (2014) compares two samples of high school seniors from 1972 and 1992 seven years after graduation and concludes that the returns to high school math score, as a measure for cognitive skills, have *increased* across these two cohorts. Although the two cohorts analyzed by Weinberger do not overlap perfectly with the two NLSY cohorts, this finding stands in direct contrast to the documented "decline" in the returns to AFQT. Weinberger's finding has received much less attention than that in the other two papers, but the contradictory findings do provide additional motivation for our analysis.

[4]See Appendix for an introduction of the background of the ASVAB, its relationship with the NLSY cohorts, as well as other essential details about the AFQT score.

characteristics of the two NLSY cohorts. In order to compare AFQT test scores across the cohorts, they created adjusted AFQT scores for the NLSY–97 respondents. Creating the adjusted scores was done in two steps. First, because the item-specific responses to the computer-based ASVAB for the NLSY–97 respondents are the property of the military, Altonji et al. relied on a concordance created by Segall (1997) that translates the computer-based IRT test scores (also known as "thetas") for the NLSY–97 cohort into paper-based scores. Altonji et al. then took Segall's paper-based AFQT scores for just the 16-year-olds in both cohorts who took the exams, and assuming rank invariance in test scores across test-taking age, assigned to respondents of different test-taking ages the AFQT test scores of 16-year-olds of the same rank. So, for example, the adjusted AFQT scores assigned to a 14-year-old test-taker with the median score in that age group is assigned to be the median AFQT of the 16-year-olds in that same NLSY cohort. We refer to these Altonji et al. adjusted AFQT scores as AAFQT scores, to make it clear that the adjustment has occurred.

It is these AAFQT scores that Altonji et al. use in their cross-cohort comparison of the NLSY cohorts. When Altonji et al. wrote their paper, the NLSY–97 cohort was still mostly in very early adulthood, and so while it was too soon to meaningfully and directly compare labor market outcomes of the two cohorts, Altonji et al. carefully documented how the observed characteristics of the two cohorts during their youth had changed, and considered the implications of those changes on wage inequality. One of the important changes Altonji et al. document is that there was a marked shift in the AAFQT distribution across the two cohorts.

Figure 4.4 plots the distributions of AAFQT scores for just the white non-Hispanic men in the two NLSY cohorts. (We limit our analysis to white non-Hispanic men in the core samples to abstract from compositional changes in the population and potential differential measurement issues across demographic groups.) The first notable finding, as highlighted

by Altonji et al., is that both the mean and the median of AAFQT scores are higher in the younger (NLSY–97) cohort. Perhaps more strikingly, the skewness of the distribution is much more pronounced for the younger cohort (0.91) than for the older cohort (0.54), with fewer scores for the younger cohort falling in the low-ish range of 125 to 150 (or so). [5]

In a later section, we return in detail to the question of whether cognitive ability has truly changed across the cohorts, but for now we treat the changing distribution of AAFQT as measuring the true change in cognitive ability. And we demonstrate that the changing distribution has important consequences for understanding the changing estimated returns to cognitive skill, as measured by AAFQT.

4.4 The Wage Return to Cognitive Skill

Many studies have used some version of NLSY AFQT test scores to estimate the labor market returns to cognitive skills. Castex and Dechter (2014) and Deming (2017) are the first studies to use AAFQT scores to measure the *changing* returns to cognitive skills across cohorts.

In Table 4.1, we report the results of regressions that specify and parameterize Equations (4.1) and (4.3) similar to those used in Castex and Dechter (2014) and Deming (2017). In particular, we estimate equations where we the labor market outcome of interest is the log of a wage measure, and we estimate a log-linear relationship between wages and AAFQT.

$$\ln W_i^t = \alpha^t + \beta^t AAFQT_i^t + \gamma^t X_i^t + \varepsilon_i^t$$

where W_i^t is an the average hourly wage of individual i from cohort t observed between the

[5] This divergence across cohorts is similarly observed among non-Hispanic white men who took the test at age 16 (see Appendix Figure 4A.1). A similar divergence is also observed for other demographic groups, including white non-Hispanic women, and indeed it is consistent with Altonji et al.'s finding for AFQT scores when pooling together all genders, races, and ethnicities (see Appendix Figures 4A.3 and 4A.2).

ages of 25 and 33.[6]

The results in Table 4.1 confirm the basic results of Castex and Dechter and Deming. Columns (1) and (4) of Table 4.1 report estimates of Equation (4.9) for the NLSY–79 and NLSY–97 cohorts respectively. The estimated (log) wage return to an additional AFQT point falls from 0.53 for the NLSY–79 cohort to 0.36 for the NLSY–97 cohort, a large and statistically significant drop of 0.17, or 47 percent. In columns (2) and (6) we add to the specification years of schooling. Including years of schooling in the specification reduces the estimated coefficients on AFQT, but the differential between the estimated coefficients across the cohorts is actually slightly larger at -0.22, representing an even larger increase in percentage terms. Adding controls for the measures of non-cognitive and social skills that Deming used, as reported in columns (3) and (6), has little additional effect on the estimated returns to AFQT. [7]

In the remainder of the paper, we focus exclusively on the univariate relationship between AAFQT and the log of wages, excluding from our analysis consideration of years of schooling or other dimensions of skill. Columns (1) and (4) of Table 4.1 thus form our baseline. We do this for four basic reasons. First, the finding that OLS wage returns to AAFQT declined across the cohorts is robust to the inclusion of those other variables. Second, years of schooling may be an mediating factor between AAFQT and labor market outcomes (Neal and Johnson 1996; Heckman, 1998). Third, to the extent that AAFQT may be mismeasured, the same can be said of the other skills measures, and considering measurement issues in multiple variables is very complex. Fourth, we examine nonparametric relationships between AAFQT and log wages, and in doing so we need to maintain

[6]Following Deming, we limit our sample in both cohorts to labor market outcomes between the ages of 25 and 33. We report results from weighted regressions, use the custom sample weights created by the BLS to keep the samples nationally representative.

[7]The finding of declining estimated wage returns to AFQT in Table 4.1 is robust to measuring wages for men of the same calendar age, it is robust to dropping calendar years when the national unemployment rate was over eight percent, and it is robust to using median regression, including zeros for years with no-employment.

a univariate framework.

In the next section, we turn to a detailed analysis of what exactly is driving the OLS estimate of the decline in the return to AAFQT across the cohorts. To do this, we hearken back to an old method for understanding how OLS estimates are constructed.

4.4.1 Implementing the Yitzhaki Decomposition

In practice, we deal with discrete X. (In particular, in our application, X is AAFQT scores.) One can discretize the construction of the OLS estimate as follows:

First, rank observations in increasing order of X, so $x_1 < x_2 < \cdots < x_n$. Let N_i be the number of duplicate observations for x_i and let N be the sum of all observations: $N = N_1 + \cdots + N_n$.

Then, let $\Delta x_i = x_{i+1} - x_i$ and let $b_i = \Delta \bar{y}_i / \Delta x_i$. Thus, we can think of b_i as the pairwise slope or estimated "unit treatment effect".

The OLS estimator can then be expressed as:

$$\beta_{OLS} = \sum_{i=1}^{n-1} w_i b_i, \quad \sum_{i=1}^{n-1} w_i = 1, \quad w_i \geq 0 \quad \forall i$$

where

$$w_i = \frac{1}{\sigma_x^2} \frac{\sum_{j=1}^{i} N_j}{N} \frac{\sum_{j=i+1}^{n} N_j}{N} \left(\frac{\sum_{j=i+1}^{n} N_j x_j}{\sum_{j=i+1}^{n} N_j} - \frac{\sum_{j=1}^{i} N_j x_j}{\sum_{j=1}^{i} N_j} \right) \Delta x_i$$

It is important to emphasize that these "Yitzhaki" weights are solely a function of X; Y's role in the construction of the OLS estimates comes only through its role in the calculation of the pairwise slopes.

Often, weighted least squares (WLS) is utilized instead of OLS. We show in the Appendix that it is straightforward to extend the Yitzhaki decomposition to WLS. Just as can

be done in weighted least squares, each observation in the sample is just "blown up" appropriately by its sample weight, and then the mechanics of the OLS Yitzhaki decomposition follows. For the remainder of the paper so as not to confuse sample weights with Yitzhaki weights, we refer to OLS throughout, but to be clear, we do apply sample weights when estimating linear regressions and when performing Yitzhaki decompositions.

Consider again the simple univariate regression equation representing the relationship between log wages and AFQT scores:

$$\ln W_i^t = \alpha^t + \beta^t AFQT_i^t + \varepsilon_i^t \tag{4.9}$$

The finding in Table 4.1, columns (1) and (4) from estimating this regression is that:

$$\beta_{OLS}^{79} > \beta_{OLS}^{97}$$

where we now have determined that

$$\beta_{OLS}^{79} = \sum_{i=1}^{n-1} w_i^{79} b_i^{79} \tag{4.10}$$

and

$$\beta_{OLS}^{97} = \sum_{i=1}^{n-1} w_i^{97} b_i^{97} \tag{4.11}$$

Thinking about the OLS estimates as weighted sums of unit treatment effects, it becomes clear that one can examine the changing OLS returns to AFQT across the NLSY–79 and NLSY–97 cohorts by examining whether the change is mechanically driven by changing Yitzhaki weights, changing pairwise slopes, or both. In particular, because the Yitzhaki weights are only a function of AAFQT scores, one can specifically examine how much the

changing distribution of AAFQT scores between the two cohorts is affecting the construction of the OLS estimates.

4.4.2 Understanding the Declining OLS Wage Returns to AAFQT

We begin by graphing in Figure 4.5 the OLS estimates of the wage returns to AAFQT in each cohort, as well as the underlying data. It is clear from Figure 4.5 that the OLS return to AAFQT is lower in the NLSY–97 cohort than in the NLSY–79 cohort, and that the result does not appear to be driven by any particular outliers.

In Figure 4.6, we graph the pairwise slopes of the Yitzhaki decomposition. These pairwise slopes come from taking the difference in the average log wages between adjacent AAFQT scores from Figure 4.5. In order to make the graph more readable, we collapse AAFQT scores into bins that each contain 3 consecutive AAFQT points, and we report the average of the pairwise slopes in each 3-point bin. One can see that it is difficult to discern any particular pattern across the AAFQT distribution or between the cohorts. This is not problematic–the pairwise slopes themselves presumably consist of a lot of random variation in wages across individuals.

The more important component of the Yitzhaki decomposition to consider is the Yitzhaki weights. In Figure 4.7, therefore, we graph the Yitzhaki weights for each NLSY cohort separately, again calculating and reporting average weights using bin sizes of 3 AAFQT points. Note that the Yitzhaki decomposition tells us that if we multiple each point in Figure 4.6 by its corresponding point in Figure 4.7, and then sum these, we recover the OLS estimate.

In Figure 4.8 we present a version of the weights that is smoothed with local linear regression. It is clear (especially in the smoothed graph in Figure 4.8) that the Yitzhaki weights differ between the two cohorts, and in particular that lower AAFQT scores receive more weight in the construction of the OLS estimate for the NLSY–97 cohort than for the

NLSY–79 cohort. It is worth considering carefully why this is happening.

Looking back at Figure 4.4, and noting the construction of the Yitzhaki weights in Equation (4.8), the increased left-skewness of the AAFQT distribution in the NLSY—97 yields larger weights on low AAFQT scores for the NLSY–97 cohort than for the NLSY79 cohort. This is true for very low AAFQT scores, but it is even true for values of the AAFQT score in the mid-low range where the density of AAFQT scores is lower in the NLSY–97 cohort.

But just seeing that the Yitzhaki weights are left-shifted in the NLSY–97 relative to the NLSY–79 does not alone explain why, mechanically, the estimated OLS returns to AAFQT are lower in the NLSY–97. To understand this, one needs to understand how the Yitzhaki weights work together with the pairwise slopes. To get an intuitive sense of this, in Figure 4.9, we again graph the Yitzhaki weights, but we also include for each cohort the pairwise slopes, but smoothed with local linear regression.

The local linear regression results reveal that the gradient of the relationship between AAFQT and log wages remained relatively upward sloping and constant through much of the AAFQT distribution for the NLSY–79 cohort (except perhaps at the very top and very bottom–places where local linear regression performs less well and where there are few obsevrations), leading to what appears to be a fairly linear and positive relationship between AAFQT and log wages for the NLSY–79 cohort. In contrast, the gradient for the NLSY–97 cohort appears much less constant; in particular, it has flat spots at various points in the distribution, especially for AAFQT scores between around 100 and 150 (which is approximately the median score in NLSY–97), and then again between around 170 and 190 (in the range of the 70th to 80th percentiles). Note that these regions of AAFQT scores are also regions that receive a lot of weight for the NLSY–97 cohort in the Yitzhaki representation of the OLS estimate, providing suggestive evidence that the lower OLS estimate of the wage return to AAFQT in the NLSY–97 cohort is being driven mainly by these flat spots in the

local linear regression.

To better understand how the different parts of the AAFQT distribution contribute to the OLS estimates of the wage returns to AAFQT, in Figure 4.10, we graph the progressive sum of the Yitzhaki equation from Equations (4.10) and (4.11), starting with the lowest AAFQT score until the entire sum is calculated and we recover the OLS estimate. That is, for each AAFQT score x_k from x_1 to x_{n-1}, we calculate for each of the two NLSY cohorts t the cumulative sum of the Yitzhaki decomposition:

$$\sum_{i=1}^{k} w_i^t b_i^t, \quad k = 1, \ldots, n-1 \tag{4.12}$$

and graph the result.

Although the results in Figure 4.10 are order-dependent so that, for example, starting from the highest AAFQT score and working out way down to the lowest would yield a different figure, we think the figure is still informative. In particular, the figure shows that the contributions of the very lowest AAFQT scores to the OLS estimates were not very different in the two cohorts, but that the progressive sums from the Yitzhaki decomposition begin to permanently diverge at an AAFQT score of around 125 (less than the 10th percentile), at which point the Yitzhaki sum for the NLSY–79 cohort rises quickly and continuously until it reaches its final level of 0.053. In contrast, the the Yitzhaki sum for the NLSY–97 cohort stays virtually constant until around an AAFQT score of 150 (the 25th percentile), then it rises quickly, only to actually fall again before recovering and rising to its final level of 0.036. This is (not surprisingly) consistent with Figure 4.9, where the flat portion of the local linear regression for the NLSY–97 cohort plays a large role in depressing the NLSY–97 OLS estimate relative to the NLSY–79 estimate.

4.5 Counterfactual OLS Estimates

Given the different OLS estimates of the wage returns to AAFQT for the NLSY–79 and NLSY–97 cohorts, we can ask the following counterfactual question: Would the estimated OLS returns to AAFQT have changed between the two cohorts if the distribution of AAFQT had not changed?

To answer this counterfactual question, we reframe it in terms of the mechanics of the Yitzhaki decomposition by equivalently asking: Would the the estimated OLS returns to AAFQT have changed between the two cohorts if the Yitzhaki weights had been held fixed across the cohorts but the observed pairwise slopes had still been realized? [8]

We answer this question by decomposing the observed difference between β_{OLS}^{79} and β_{OLS}^{97} into two ways:

$$\begin{aligned} \beta_{OLS}^{79} - \beta_{OLS}^{97} &= \left(\beta_{OLS}^{79} - \beta_{OLS}^{97}|w^{79}\right) + \left(\beta_{OLS}^{97}|w^{79} - \beta_{OLS}^{97}\right) \\ &= \left(\beta_{OLS}^{79} - \beta_{OLS}^{79}|w^{97}\right) + \left(\beta_{OLS}^{79}|w^{97} - \beta_{OLS}^{97}\right) \end{aligned} \tag{4.13}$$

The first term in parentheses in the top decomposition is the counterfactual difference in the OLS estimates, holding fixed the AAFQT distribution (and the corresponding Yitzhaki weights) at the NLSY–79 level. Graphically, the components of this counterfactual decomposition can be represented by Figure 4.11, where we overlay the Yitzhaki weights from NLSY–79 onto the (smoothed representation) of the observed pairwise slopes from the two cohorts.

Alternatively, the second term in parentheses in the bottom decomposition is the counterfactual difference in the OLS estimates, holding fixed the AAFQT distribution (and the corresponding Yitzhaki weights) at the NLSY–97 level. Graphically, the components of

[8] We believe we are the first to use the Yitzhaki decomposition to consider this kind of counterfactual comparison across cohorts.

this counterfactual decomposition can be represented by Figure 4.11, where here we overlay the Yitzhaki weights from NLSY–97 onto the (smoothed representation) of the observed pairwise slopes from the two cohorts.

As a reminder, the actual OLS estimates in Table 4.1 Columns (1) and (4) demonstrate that the estimated wage return to AAFQT declined meaningfully (by 37%) across the cohorts. Formally:

$$\beta_{OLS}^{79} - \beta_{OLS}^{97} = 0.54 - 0.34 = 0.20$$

Using the top decomposition in Equation (4.13), this decline can be decomposed as:

$$\begin{aligned}
\beta_{OLS}^{79} - \beta_{OLS}^{97} &= \sum_i w_i^{79}(b_i^{79} - b_i^{97}) + \sum_i (w_i^{79} - w_i^{97}) b_i^{97} \\
&= \left(\beta_{OLS}^{79} - \beta_{OLS}^{97} | w^{79}\right) + \left(\beta_{OLS}^{97} | w^{79} - \beta_{OLS}^{97}\right) \\
&= \left(0.53 - 0.78\right) + \left(0.78 - 0.36\right) \\
&= -0.25 + 0.42
\end{aligned}$$

The term in the first set of parentheses is the counterfactual change in returns holding the Yitzhaki weights at NLSY–79 levels. This counterfactual indicates that the OLS return to AAFQT scores actually went up by 0.25 points, a substantial *increase* of 47%.

Alternatively, if we fix the Yitzhaki weights at NLSY–97 levels, we find this:

$$\begin{aligned}
\beta_{OLS}^{79} - \beta_{OLS}^{97} &= \sum_i (w_i^{79} - w_i^{97}) b_i^{79} + \sum_i w_i^{97}(b_i^{79} - b_i^{97}) \\
&= \left(\beta_{OLS}^{79} - \beta_{OLS}^{79} | w^{97}\right) + \left(\beta_{OLS}^{79} | w^{97} - \beta_{OLS}^{97}\right) \\
&= \left(0.53 - 0.61\right) + \left(0.61 - 0.36\right) \\
&= -0.08 + 0.25
\end{aligned}$$

The term in the second set of parentheses above is the counterfactual change in returns holding weights at NLSY–97 levels. Hence, when using NLSY–97 weights, the estimated return to AAFQT falls by 0.25 points, somewhat larger even than the observed OLS estimate.

This finding highlights the critical role that the changing composition of AAFQT scores across the two cohorts has played in the narratives that there has been a decline in the return to cognitive ability.

We note that Castex and Dechter (2014) use AAFQT scores directly in their analysis, as we have done here. Deming (2017), on the other hand, does not. He standardizes AAFQT's in both cohorts to have a mean of zero and a standard deviation of one. It should be clear from the analysis we have done here that standardizing and equating the distributions across the two cohorts is not a benign transformation–in particular, it sets the Yitzhaki weights to be the the same across the two cohorts in a context where the weights end up playing a key role in determining the OLS estimates.

Although in Section 4.2 we developed a model to motivate the changing distribution of cognitive skill and its effect on wage returns, in the next section we delve into whether we can really be sure that the returns to cognitive skill have changed. In particular, we consider the role of measurement error.

4.6 The Measurement of Cognitive Ability

Consideration of the economic underpinnings of rising returns to cognitive ability rests on the assumption that estimated rising returns actually reflect reality, and are not a result of issues in the measurement of cognitive ability.

There has long been concern about whether test scores correctly measure the underlying skills they are meant to capture, and there have been various proposed econometric

approaches to correcting for measurement error (see, e.g., Griliches and W. Mason, 1972, for an early treatment of this issue). In this section, we delve into three related questions: (1) whether AAFQT is a mismeasured proxy for cognitive skill; (2) if so, at what stage much of the measurement error may have been introduced; (3) and whether it is possible to easily correct for any measurement error.[9]

Our substantive interest remains focused on whether or not the wage return to cognitive skill has changed. Given our previous finding that the answer to this question is strongly dependent on the changing distribution of AAFQT across the cohorts, and given that we have no true measure of cognitive skill for either cohort, we emphasize measurement issues within the context of the *changing* distributions of AAFQT across the NLSY cohorts.[10]

The first thing to note about the increased (left) skewness of AAFQT scores in the NLSY–97 relative to those for the earlier cohort is that, as we show in Appendix Figure 4A.4, it is not being driven by one specific section of the adjusted ASVAB scores (adjusted by Segall to turn the IRT-based scores into paper-based scores) that Altonji et al. used to create AAFQT. Indeed, it is especially pronounced in both Numerical Operations and Paragraph Comprehension–both of which also are parts of the AFQT and may well reflect different underlying dimensions of cognitive ability.

Second, this divergence in AAFQT scores across the cohorts is not a function of Altonji et al.'s age adjustment. This has to be the case if Altonji et al. used the rank order of scores for non-16-year-olds and assigned by rank the numerical scores of 16-year-olds. We confirm this in in Appendix Figure 4A.1 where we graph the AAFQT scores for just 16-

[9]There is a separate but important measurement debate about whether the AFQT test is racially biased (Wigdor and Green Jr., 1991; Neal and Johnson 1996; Rodgers and Spriggs, 1996; Heckman, 1998). We limit our analysis to non-Hispanic whites.

[10]Recall that the concordance of the AFQT score across cohorts that was done by Altonji et al. involves two critical steps. The first step is a mapping between the computer-based test scores (in the NLSY–97) and paper-based test scores (in the NLSY–79), which relies on the external study conducted by Segall (1997) for the Department of Defense. The second step is a mapping between test scores at age 16 and test scores at other ages. As discussed above, the second step relies on a somewhat strong assumption that the AFQT score is rank-invariant across ages.

year-olds in our sample, and in Appendix Figure 4A.5 where we also graph each section of the adjusted ASVAB scores for 16-year-olds as well.

Third, the divergence does not seem to be a direct artifact of Segall's conversion of computer-based IRT scores to paper-based scores. In Appendix Figures 4A.6, we graph the distribution of the original IRT-based scores for different ASVAB sections for the NLSY–97 cohort alongside IRT-based scores that were constructed after-the-fact from the original paper-based tests by researchers from the Ohio State University for the NLSY–79 cohort.[11] Appendix Figure 4A.7 shows the same distribution for just 16-year-olds in our sample. The divergence of scores in the NLSY–97 IRT-based scores relative to NLSY–79 scores is clearly visible in different sections.

However, despite the existence of IRT-based scores ("thetas") for both NLSY cohorts, it far from clear that they are directly comparable for at least three reasons. First, we do not know if the IRT models and estimation methods used for the two cohorts are the same. Second, even if the models and methods are the same, the raw data imported to the models may still not be comparable due to the very different test formats. Third, there is one important hint that something is amiss in the IRT scores for the NLSY–97.

Appendix Figure 4A.8 plots the standard errors of the estimated IRT scores, the "thetas," for different ASVAB sections. As pointed out in past studies (Schofield, 2014; Jacob and Rothstein, 2016, in IRT models, thetas in IRT models are more precisely estimated for the middle of the distribution, leading to a non-classical measurement error structure with larger errors at the tails. Thus, the measurement error in the test score is correlated with the underlying true ability or skill.[12] The error structure of the thetas in the NLSY–79 is

[11] See Ing, Lunney, and Olsen (2012) for more detailed discussion. IRT-based scores ("thetas") are not available for the numerical operations section of ASVAB in the NLSY–79.

[12] This particular measurement error issue of IRT scores is a feature of IRT, generally, and not just for NLSY datasets. Jacob and Rothstein (2016) point out this issue for the many longitudinal studies conducted by National Center for Education Statistics (NCES), such as the National Education Longitudinal Study (NELS).

generally symmetric, exactly as expected from the IRT model (Schofield 2014). What is somewhat odd is that in the NLSY–97, the errors are abnormally large for low ASVAB scorers.[13] This empirical irregularity of the thetas for the later cohort is particularly worrisome given our finding of the importance of low AAFQT scores in driving the estimates of changing returns to cognitive skill.

All told, it appears that to the extent that there may be important measurement error in the AAFQT scores in the NLSY–97 that drives the divergence across the cohorts, it seems to be present in the original IRT-based AFQT test score results for the NLSY–97 respondents, and is not a function of Segall's conversion to paper-based scores or Altonji et al.'s age adjustment.

Aware of the potential bias caused by measurement errors, Castex and Dechter (2014) used SAT scores as an instrumental variable for AFQT scores in some of their analyses. In general, in order to use one test score as an instrument for another, one must assume that measurement errors in the two test scores are both uncorrelated with the true underlying skill and uncorrelated with each other. In the case of the IRT-based AFQT scores in the NLSY–97, the measurement error is non-classical by the nature of the IRT process itself, so as Schofield (2014) points out, the simple IV strategy used by Castex and Dechter (2015) does not work. One alternative approach proposed by Junker, Schofield, and Taylor (2012) is a "mixed effects structural equations" (MESE) method, which jointly estimates the wage equation and the IRT model. In this way, the errors in the IRT estimation process can be incorporated into the estimation of the wage equation. To apply the MESE method in the estimation of wage returns to cognitive skills using the NLSYs, we would need the data on the actual item-level responses of the individual NLSY test-takers to each question in the ASVAB. These item-level responses are currently available for the NLSY–79 but not for

[13] We thank Dan Black for pointing this out to us.

the NLSY–97. [14]

In the end, we do not have enough information with which to assess the true extent to which measurement problems in the AFQT exist. Moreover, we cannot correct for them in the estimation of the (changing) return to cognitive skill. We do have enough information, however, to be concerned that the changing distribution of AFQT scores (and thus, AAFQT scores), may be an artifact of measurement error.

4.7 Conclusion

From an underlying economic perspective, if one wants to specify a Mincer-style OLS regression of log wages on cognitive skill (as proxied by AAFQT scores), where returns to cognitive skill are changing over time, one needs to take a stand on why these returns are changing. Once the (reasonable) assumption is made that these changes over time are being driven by technological change in the way that skills are used in firms, one needs to then consider how endogenous investments in cognitive skill accumulation will change. We show in a simple investment model that if the marginal productivity of cognitive skill decreases due to technological change, both cognitive skill investments and labor market returns to cognitive investment will decrease.

By reviving the Yitzhaki Decomposition, an old method for understanding the construction of least squares estimates, we demonstrate that declining estimated returns to AAFQT scores in the NLSY cohorts are driven by changes in the distribution of AAFQT scores between the two cohorts, and in particular by the increased left skewness of the distribution

[14] Issues in the measurement of labor market outcomes could also lead to biased estimates of the returns to cognitive ability. In results available upon request, we show that the declining estimated wage return to AAFQT in the NLSY are robust to reasonable alternatives. In particular, we find declining returns when we use annual measures of hourly wages rather than averaging wage data across multiple years, when using median instead of mean regression; when we exclude years when the national unemployment rate was greater than 8.0; and when we account for the labor force attachment of individuals throughout the sample periods.

of AAFQT in the NLSY–97. To the extent that returns to cognitive ability did decline for the white men in our sample, the Yitzhaki weights and the accompanying unit slopes show that they did so only at the lower part of the AAFQT distribution, where the distribution of AAFQT scores was markedly changing across the cohorts. Rationalizing this within the context of an economic model of cognitive skill investment in the face of changing technological skill requires a more complex model than we have outlined, and is a goal of future research.

Nonetheless, we want to end with a cautionary note. While it is tempting for economists to assume that our economic models are correct when they fit unusual patterns we see in data, we need to be humble enough to consider the possibility that our models are only as good as the data with which we test them. As we have shown, the conclusion that cognitive skills in the U.S. labor market declined between the two NLSY cohorts relies on the assumption that the shift in the observed distribution of AAFQT scores across the cohorts is real (and that we have correctly specified the relationship between wages and cognitive skill). We have shown that the shift in AAFQT scores may be an artifact of measurement issues in the recorded and unadjusted AFQT scores for the two cohorts, and especially for the NLSY-97 cohort. We cannot prove this, and we cannot correct for it given the data that we have. Thus, the question of whether the return to cognitive skills has declined in the U.S. remains an open one.

Figure 4.1: Simulated distribution of H_t before (period 1) and after (period 2) increase in productivity

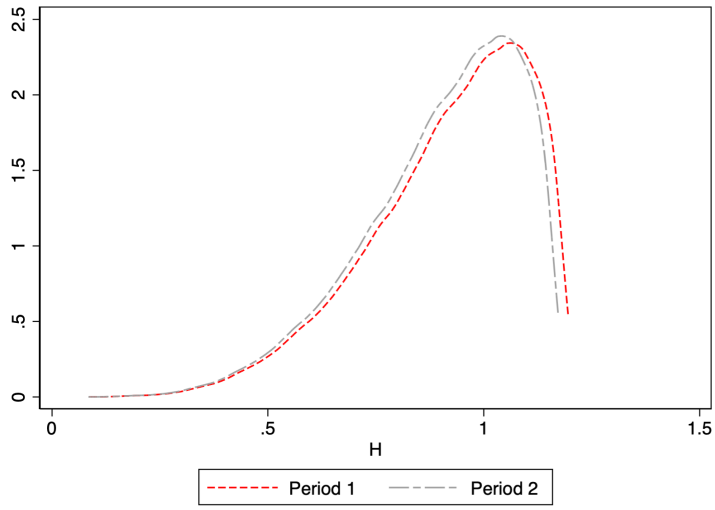


Figure 4.2: Simulated Yitzhaki slopes $b(h)$ before (period 1) and after (period 2) increase in productivity

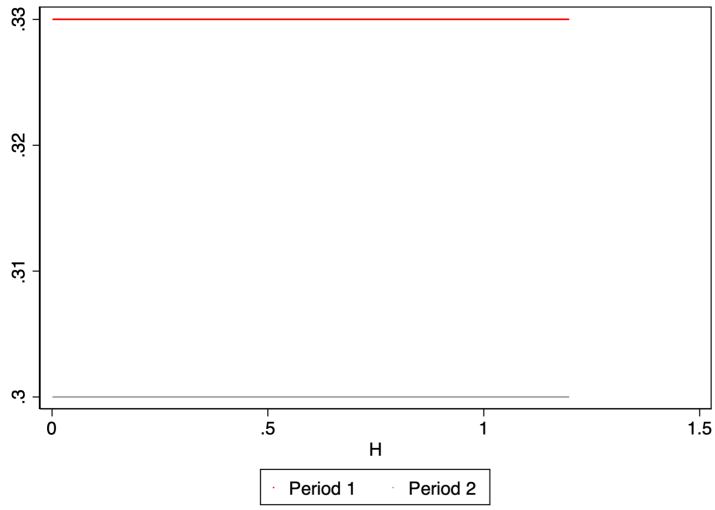


Figure 4.3: Simulated Yitzhaki weights $w(h)$ before (period 1) and after (period 2) increase in productivity

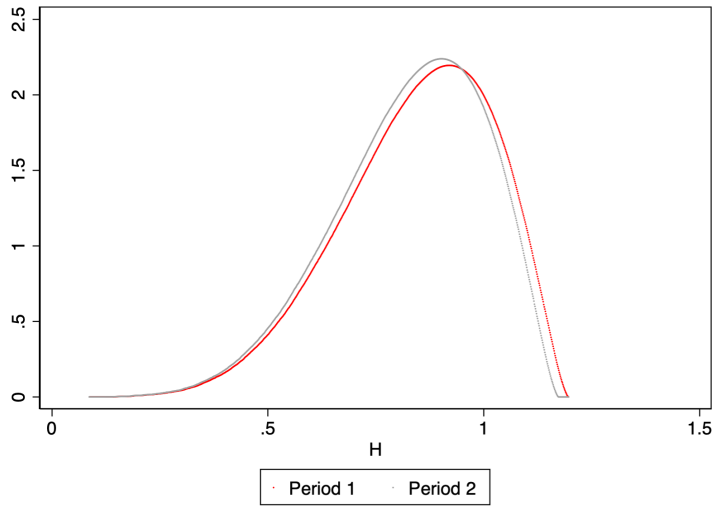


Figure 4.4: Adjusted AFQT test scores for White Non-Hispanic Men

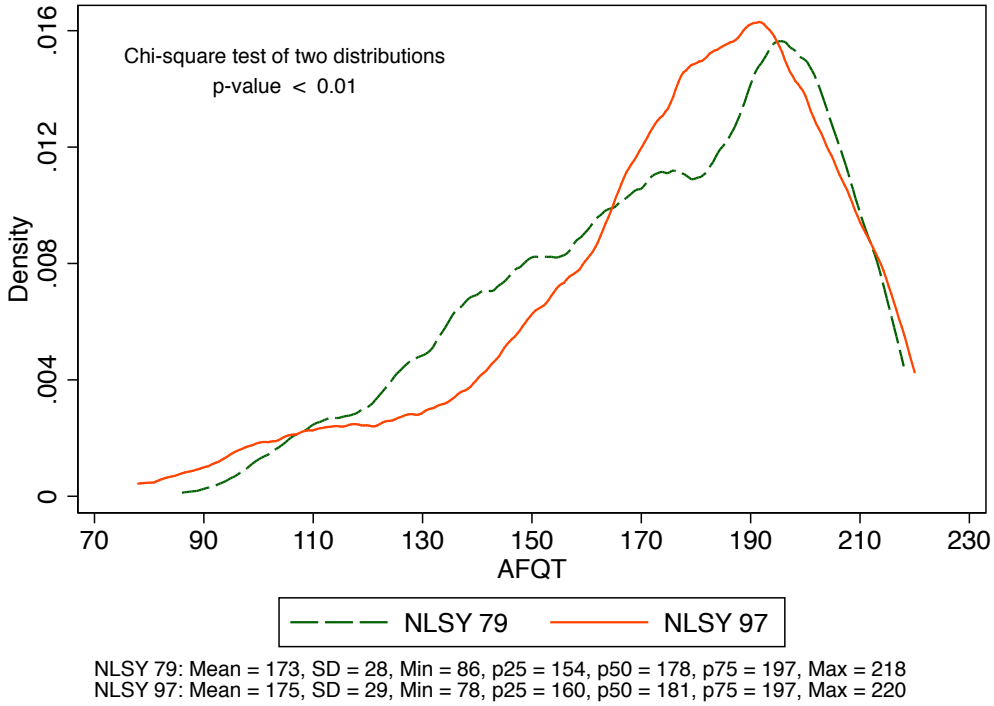


Figure 4.5: OLS Estimates of the Wage Returns to AAFQT

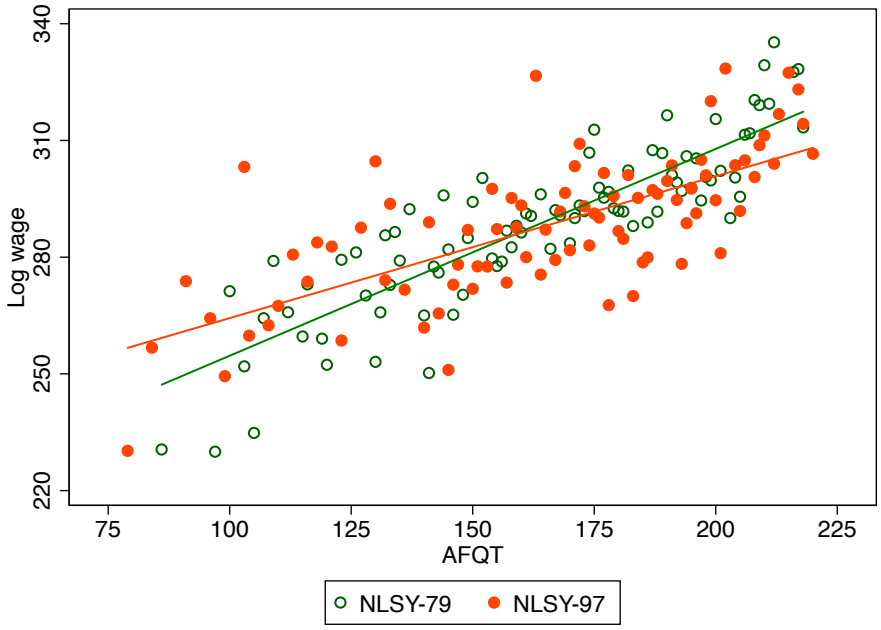
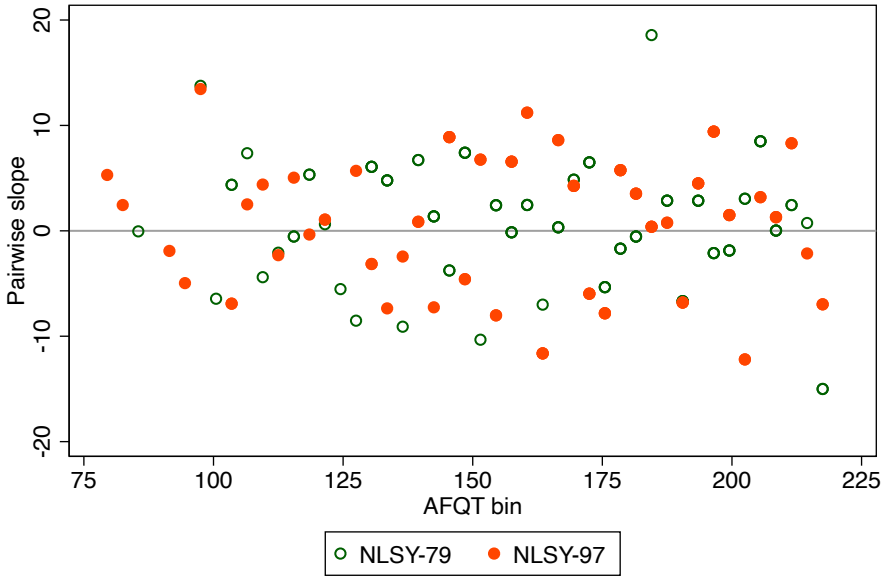


Figure 4.6: Pairwise Slopes



Pairwise slopes in 79: Mean = 1.221, SD = 9.92, OLS = .531
Pairwise slopes in 97: Mean = .472, SD = 15.584, OLS = .364

Figure 4.7: Yitzhaki Weights

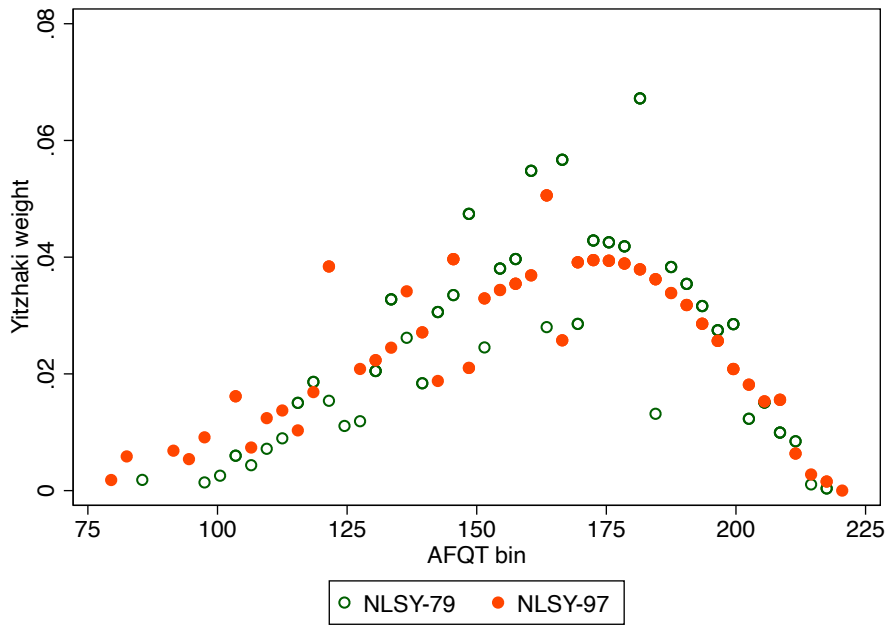
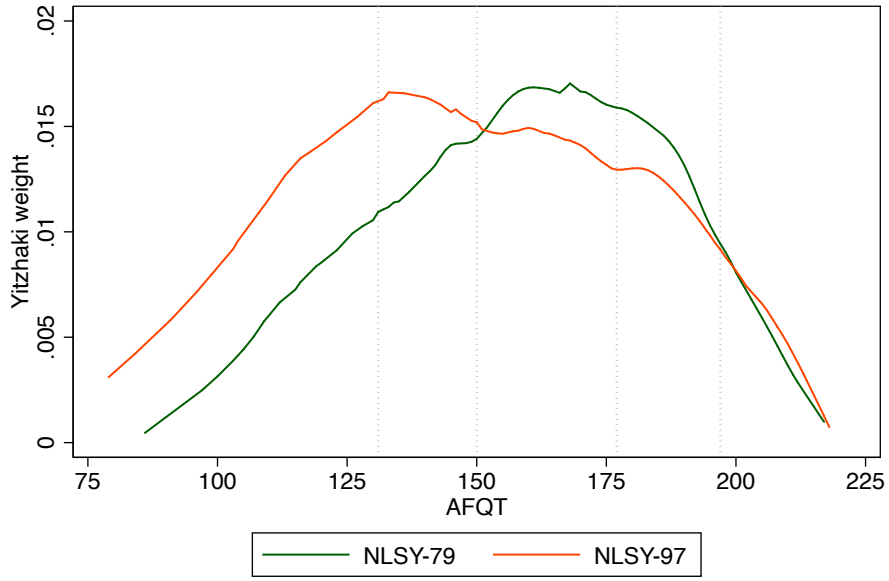
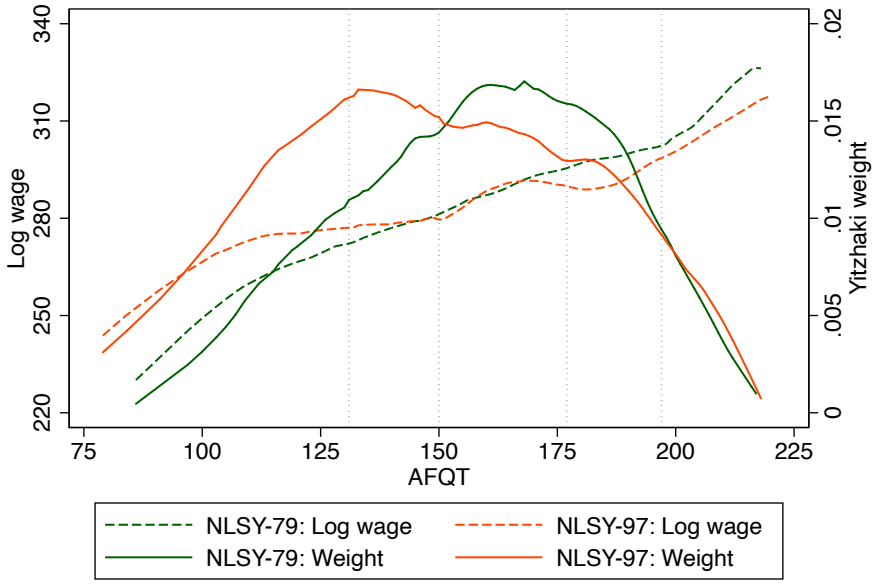


Figure 4.8: Smoothed Yitzhaki Weights



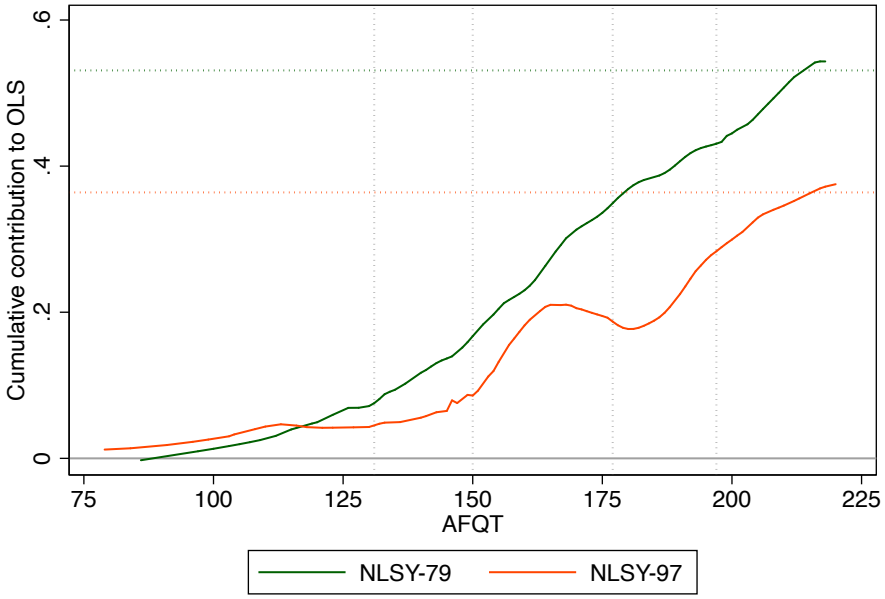
Lowess: bandwidth of 0.3
The vertical lines are p10, p25, p50, and p75 of NLSY-79

Figure 4.9: Smoothed Yitzhaki Weights and Local Linear Regression



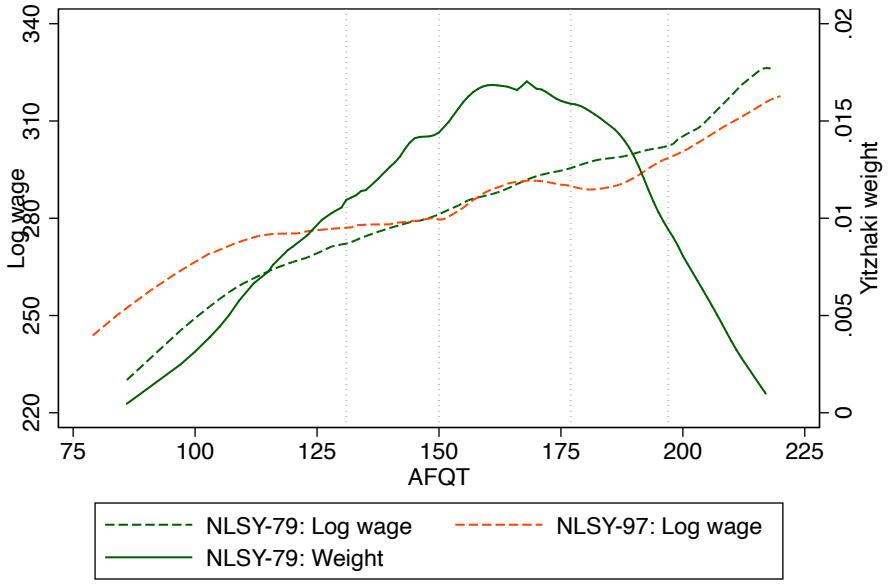
Lowess: bandwidth of 0.3
The vertical lines are p10, p25, p50, and p75 of NLSY-79

Figure 4.10: Yitzhaki Weights: Cumulative Contributions



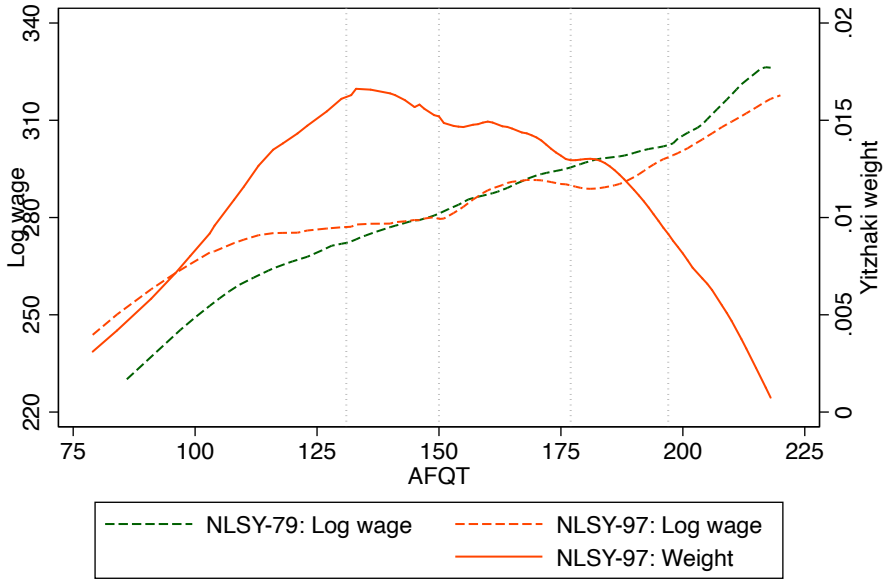
Lowess: bandwidth of 0.3; OLS in 79 = .531, OLS in 97 = .364

Figure 4.11: Weights Fixed at NLSY–79 Levels



Lowess: bandwidth of 0.3
The vertical lines are 10th pctile (131), 25th pctile (150), median (177), and 75th pctile (197) of NLSY-7

Figure 4.12: Weights Fixed at NLSY–97 Levels



Lowess: bandwidth of 0.3
The vertical lines are 10th pctile (131), 25th pctile (150), median (177), and 75th pctile (197) of NLSY-7

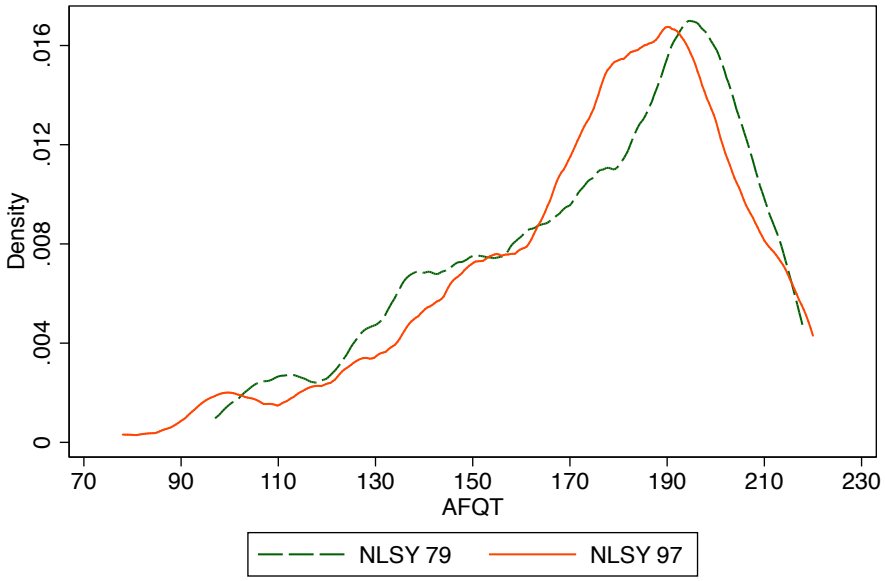
Table 4.1: OLS estimates

	NLSY–79			NLSY–97		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
AFQT	0.53**	0.37**	0.32**	0.36**	0.15**	0.16**
	[0.03]	[0.04]	[0.04]	[0.04]	[0.05]	[0.05]
Education		2.97**	2.63**		4.18**	3.79**
		[0.44]	[0.44]		[0.48]	[0.48]
Change of AFQT from NLSY–79				–0.17	–0.22	-0.16
P-value				(0.00)	(0.00)	(0.01)
Change of Education from NLSY–79					1.20	1.16
P-value					(0.07)	(0.08)
Social and non-cog scores			X			X
Observations	2,080	2,080	2,080	1,565	1,565	1,565
Adjusted R-squared	0.12	0.14	0.16	0.04	0.09	0.12

Note: ** $p < 0.01$; NLS custom sample weights are used.

Appendix Figures

Figure 4A.1: Adjusted AFQT test scores for White Non-Hispanic Men who took the test at age 16



NLSY 79: Mean = 174, SD = 29, Min = 97, p25 = 154, p50 = 180, p75 = 197, Max = 218
NLSY 97: Mean = 174, SD = 29, Min = 78, p25 = 158, p50 = 180, p75 = 195, Max = 220

Figure 4A.2: Adjusted AFQT test scores at age 16 pooling all genders, races and ethnicities (Altonji et al. 2012)

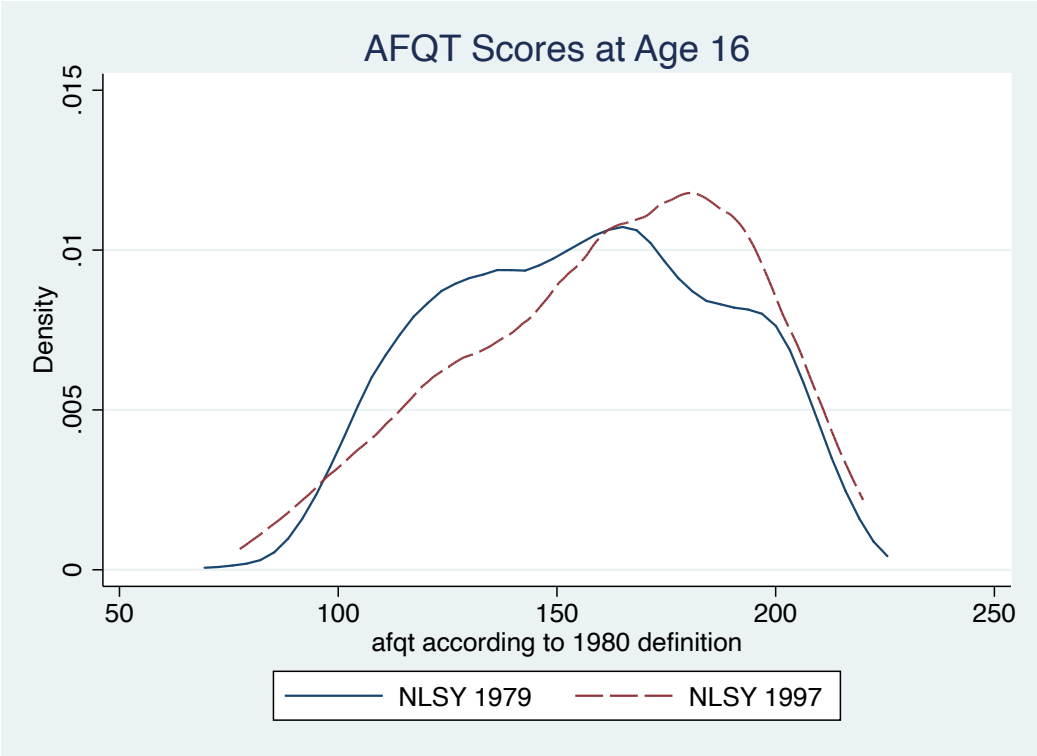


Figure 4A.3: Adjusted AFQT test scores for White Non-Hispanic Men and Women

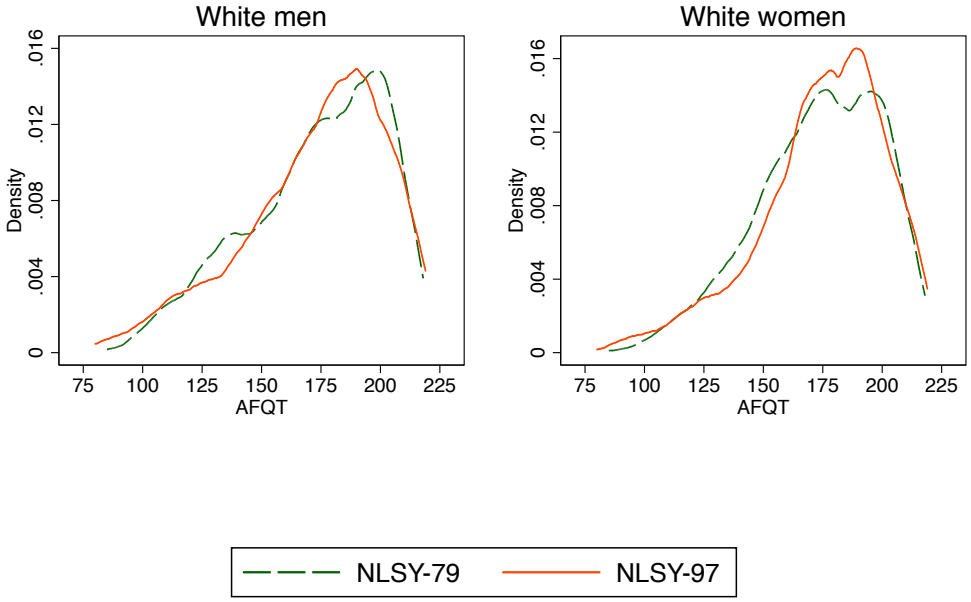
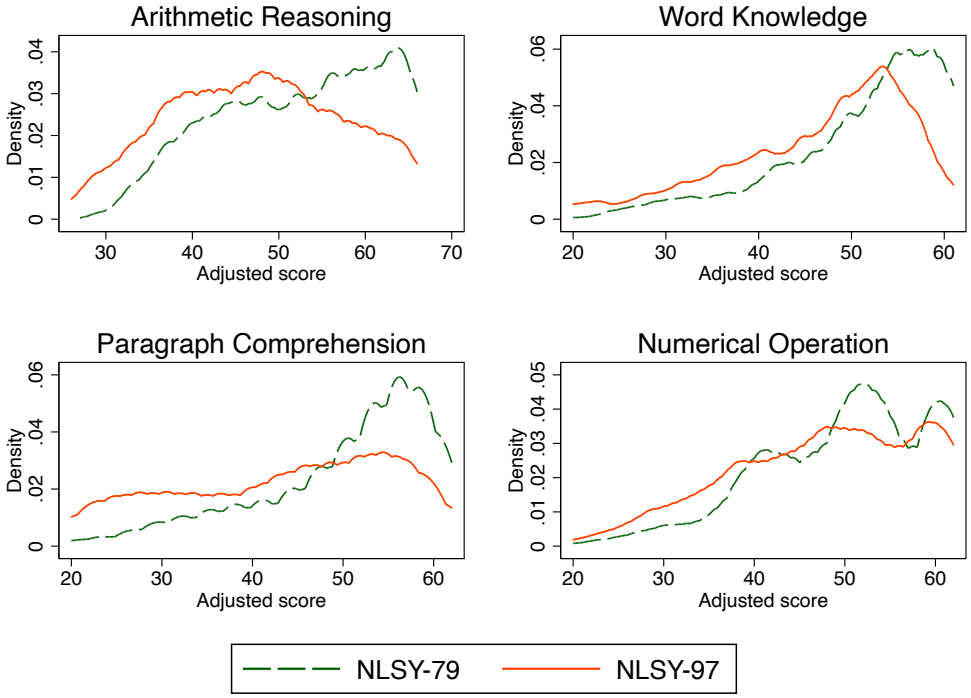
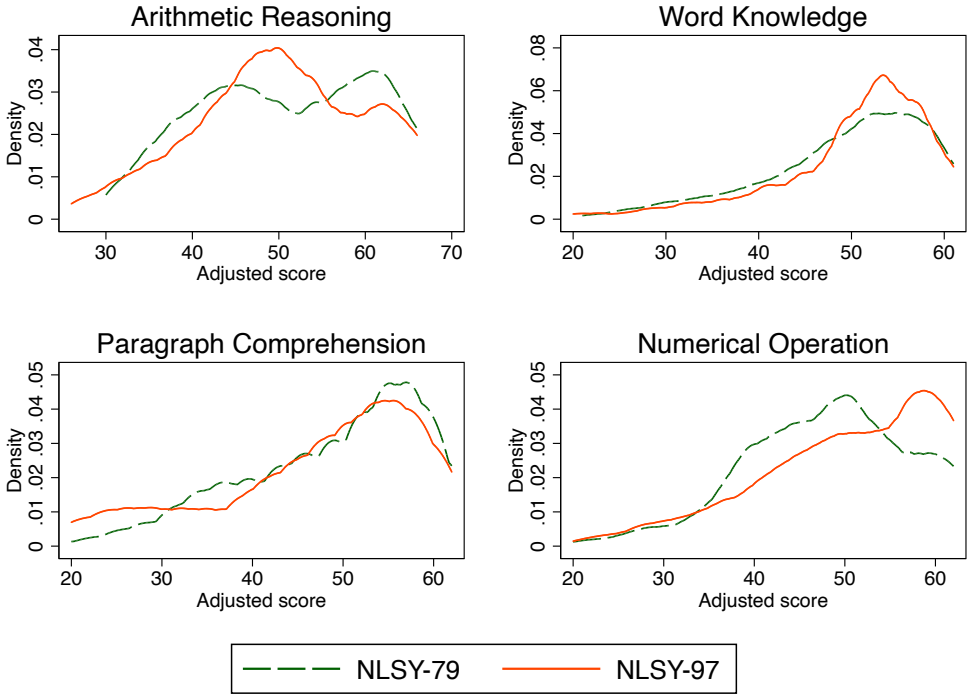


Figure 4A.4: Adjusted ASVAB subsection scores for White Non-Hispanic Men



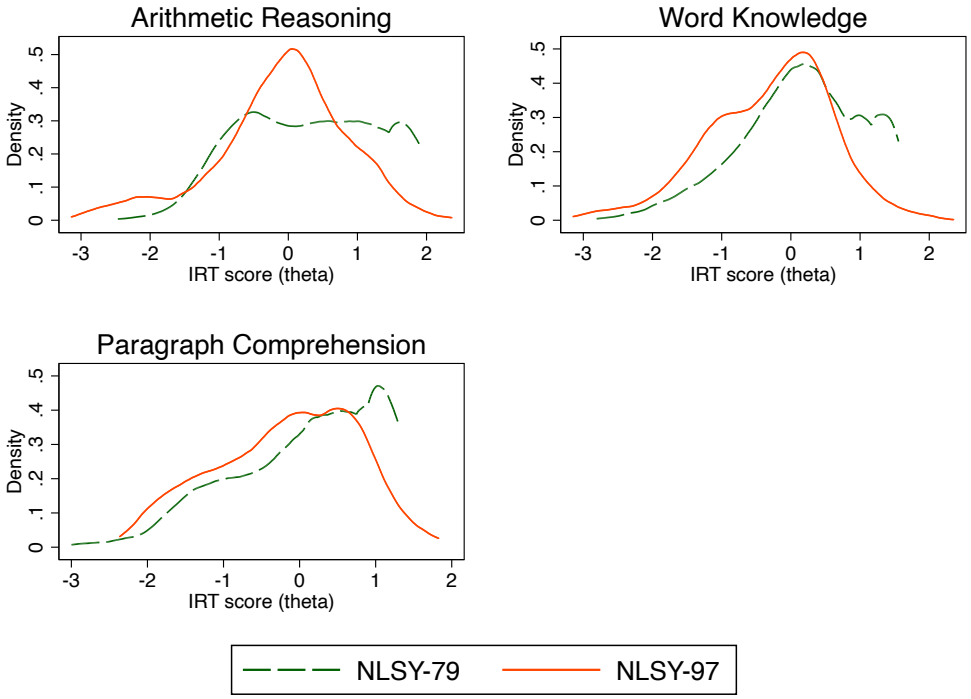
Notes: The AFQT score is constructed based on four ASVAB sections: arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operation.

Figure 4A.5: Adjusted ASVAB subsection scores for White Non-Hispanic Men at age 16



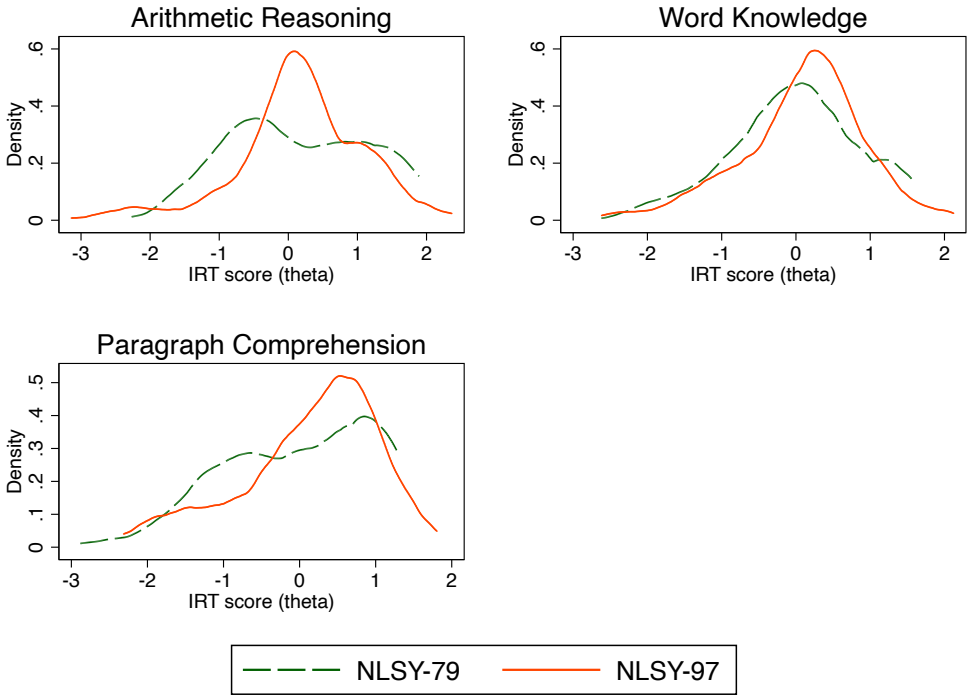
Notes: The AFQT score is constructed based on four ASVAB sections: arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operation.

Figure 4A.6: IRT-based ASVAB subsection scores for White Non-Hispanic Men



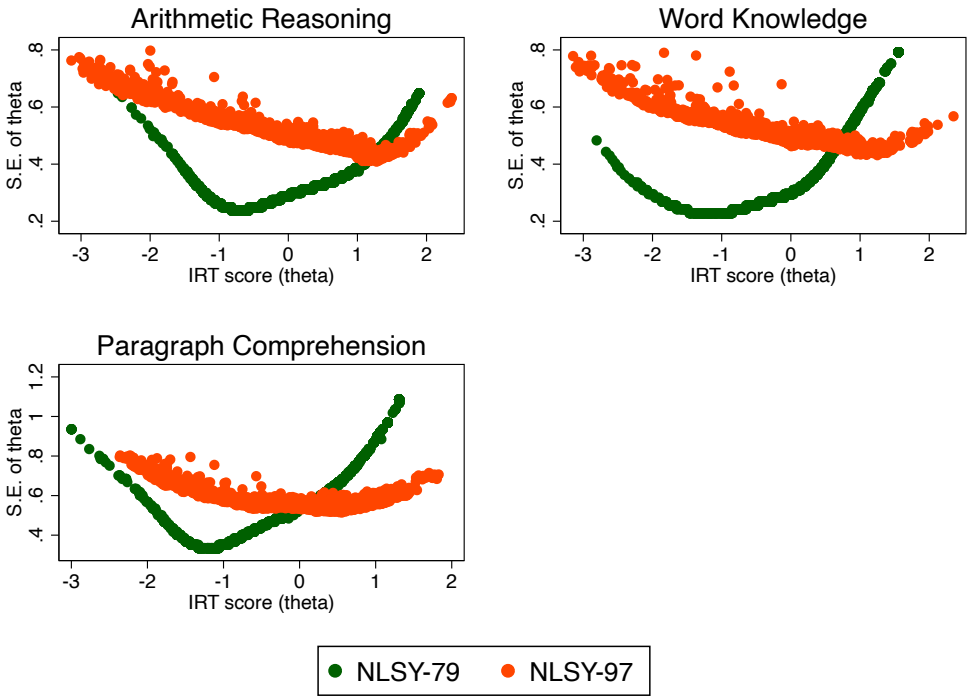
Notes: The AFQT score is constructed based on four ASVAB sections: arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operation. IRT-based scores are not available for numerical operation in the NLSY–79.

Figure 4A.7: IRT-based ASVAB subsection scores for White Non-Hispanic Men at age 16



Notes: The AFQT score is constructed based on four ASVAB sections: arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operation. IRT-based scores are not available for numerical operation in the NLSY–79.

Figure 4A.8: Standard Errors of IRT-based ASVAB subsection scores for White Non-Hispanic Men



Notes: MK=math knowledge, AR=arithmetic reasoning, WK=word knowledge, PC=paragraph comprehension. The standard errors of estimated IRT-based scores ("thetas") are plotted against "thetas".

Yitzhaki Decomposition with Weights

For simplicity, Yitzhaki's decomposition formula (Proposition 1 in Yitzhaki 1996) assumes that each value of X has only one observation. In practice, each value of X can be linked to multiple observations in the data. As suggested by Yitzhaki (1996), all observations with the same X should be aggregated, leading to a *grouped* dataset in which the outcome Y is averaged within each value of X. In a univariate model, we can recap the original OLS estimate by using the grouped data and weighting the grouped regression with group size. In addition, each observation in the data can represent multiple observations in the population. It is sometimes more appropriate to use Weighted Least Squares (WLS) with sample weights than OLS (Solon, Haider, and Wooldridge 2015). In this appendix, we extend Yitzhaki's formula to allow for these two types of weights.

Following Yitzhaki's notation, let y_i and x_i $(i = 1, \ldots, n)$ be observations and ranked in an increasing order of X. An important simplification that Yitzhaki makes is that $\Delta x_i = x_{i+1} - x_i > 0$, i.e., each value of X has only one observation. Here we extend Yitzhaki's set-up and allow there to be duplicate observations. Let there be N_i duplicate observations for (x_i, y_i). Let $b_i = \Delta y_i / \Delta x_i$ be the slope of two adjacent values of X.

Like Yitzhaki (1996), we are interested in decomposing the point estimate. Given this, the two types of weights mentioned above are both equivalent to adding duplicate observations. The distinction between the two cases is the construction of y_i. In the first case (without sample weights), y_i is the average of all Y linked to x_i. In the second case (with sample weights, i.e. WLS), y_i is the weighted average of all Y linked to x_i.

With duplicate observations, the sample covariance of Y and X can be expressed as:

$$\begin{aligned}
\operatorname{cov}(y,x) &= \tfrac{1}{2n(n-1)} \textstyle\sum_{i=1}^{n} \sum_{j=1}^{n} N_i N_j \left(x_i - x_j\right)\left(y_i - y_j\right) \\
&= \tfrac{1}{n(n-1)} \textstyle\sum_{i=2}^{n} \sum_{j=1}^{i-1} N_i N_j \left(x_i - x_j\right)\left(y_i - y_j\right)
\end{aligned}$$

Note that when there are no duplicate observations ($N_i = 1, \text{for all} i$), the expression becomes $\text{cov}(y,x) = \frac{1}{n(n-1)} \sum_{i=2}^{n} \sum_{j=1}^{i-1} \left(x_i - x_j\right)\left(y_i - y_j\right)$, which is what Yitzhaki presents in Proposition 1 (Yitzhaki 1996).

Like Yitzhaki, we substitute $\left(x_i - x_j\right) = \Delta x_i + \Delta x_{i+1} + \cdots + \Delta x_{j-1}$ and $\left(y_i - y_j\right) = b_i \Delta x_i + b_{i+1}\Delta x_{i+1} + \cdots + b_{j-1}\Delta x_{j-1}$. After collecting like terms, we get:

$$\text{cov}(y,x) = \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \left\{ \sum_{j=i}^{n-1} \left(N_1 + \cdots N_i\right)\left(N_{j+1} + \cdots + N_n\right) \Delta x_j \right.$$
$$\left. + \sum_{j=1}^{i-1} \left(N_{i+1} + \cdots + N_n\right)\left(N_1 + \cdots N_j\right) \Delta x_j \right\} \Delta x_i b_i$$

Again, when there are no duplicate observations, the expression simplifies to $\text{cov}(y,x) = \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \left\{ \sum_{j=i}^{n-1} i(n-j)\Delta x_j + \sum_{j=1}^{i-1} j(n-i)\Delta x_j \right\} \Delta x_i b_i$, as in Yitzhaki (1996).

Similarly, we can get the expression for $\text{cov}(x,x)$:

$$\text{cov(x, x)} = \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \left\{ \sum_{j=i}^{n-1} \left(N_1 + \cdots N_i\right)\left(N_{j+1} + \cdots + N_n\right) \Delta x_j \right.$$
$$\left. + \sum_{j=1}^{i-1} \left(N_{i+1} + \cdots + N_n\right)\left(N_1 + \cdots N_j\right) \Delta x_j \right\} \Delta x_i$$

We can then write down the OLS/WLS estimator as a weighted average of b_i:

$$\widetilde{b_{OLS/WLS}} = \frac{\text{cov}(y,x)}{\text{cov}(x,x)} = w_i b_i, \quad \text{where } \sum_{i=1}^{n-1} w_i = 1$$

where the weight w_i is:

$$w_i = \frac{\left\{ \sum_{j=i}^{n-1} \left(N_1 + \cdots N_i\right)\left(N_{j+1} + \cdots + N_n\right) \Delta x_j + \sum_{j=1}^{i-1} \left(N_{i+1} + \cdots + N_n\right)\left(N_1 + \cdots N_j\right) \Delta x_j \right\} \Delta x_i}{\sum_{k=1}^{n-1} \left\{ \sum_{j=i}^{n-1} \left(N_1 + \cdots N_k\right)\left(N_{j+1} + \cdots + N_n\right) \Delta x_j + \sum_{j=1}^{k-1} \left(N_{k+1} + \cdots + N_n\right)\left(N_1 + \cdots N_j\right) \Delta x_j \right\} \Delta x_k}$$

The numerator of w_i can be written equivalently in a more intuitive expression:

$$\left(\sum_{j=1}^{i} N_j\right)\left(\sum_{j=i+1}^{n} N_j\right)\cdot\left(\frac{\sum_{j=i+1}^{n} N_j x_j}{\sum_{j=i+1}^{n} N_j}-\frac{\sum_{j=1}^{i} N_j x_j}{\sum_{j=1}^{i} N_j}\right)\cdot\Delta x_i$$

As a comparison, the continuous version of the weighting function $w(x)$ is:

$$w(x)=\frac{F_X(x)\cdot(1-F_X(x))}{\sigma_X^2}\{E(X\mid X>x)-E(X\mid X\leq x)\}$$

The first term in the discrete weighting function $\left(\sum_{j=1}^{i} N_j\right)\left(\sum_{j=i+1}^{n} N_j\right)$ matches with $F_X(x)\cdot(1-F_X(x))$ in the continuous weighting function. The second term in the discrete weighting function $\frac{\sum_{j=i+1}^{n} N_j x_j}{\sum_{j=i+1}^{n} N_j}-\frac{\sum_{j=1}^{l} N_j x_j}{\sum_{j=1}^{i} N_j}$ matches with $E(X\mid X>x)-E(X\mid X\leq x)$ in the continuous weighting function. Compared to the case with no duplicate observations, here both the cumulative density and the conditional expected value are expressed in a weighted form.

Appendix A: A Note on Armed Forces Qualification Test (AFQT) score

In this section, I describe the background and essential details of the Armed Forces Qualification Test (AFQT) score. This is a collection of information from different sources: a manuscript by Altonji, Bharadwaj, and Lange (2009), a technical bulletin by Defense Manpower Data Center (2006) which includes several chapters from Sands, Waters, and McBride (1997), annual reports on population representation in the military services (e.g. Quester and Shuford, 2017), and the introduction on the NLSY website (Bureau of Labor Statistics, n.d.(b); Bureau of Labor Statistics, 1992; Bureau of Labor Statistics, n.d.(a)). 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 been using the AFQT score, as well as other tests in the ASVAB, to measure skills and abilities (Neal and Johnson, 1996; Heckman, Stixrud, and Urzúa, 2006; Altonji, Bharadwaj, and Lange, 2012; Prada and Urzúa, 2017). This is facilitated by the data of the NLSY–79 and the NLSY–97, as survey respondents took the ASVAB test.

A.1 A short history of the ASVAB and the NLSY

The ASVAB was first introduced in 1968 and has undergone several adjustments and revisions since. One important adjustment has been to update the norms of the ASVAB (Defense Manpower Data Center, 2006). In practice, the military sets a goal of selecting

only applicants who rank higher than X% of American youth in the national distribution of ability and skill. Different military branches have different qualification cutoffs, and many of them now use a cutoff of 30%–40% for applicants with a high school diploma. Recruiters therefore need to know how the Xth percentile youth in the national population scores on the ASVAB in order to compare military applicants to this benchmark. To ensure that contemporary applicants are always compared to an appropriate benchmark, the benchmark must be updated over time.[1]

In 1979, after questioning the appropriateness of using the World War II reference population as the benchmark, the DOD and Congress decided to let the NLSY–79 respondents take the ASVAB, and the DOD used their scores as the new benchmark for military enlistees. The NLSY–79 served as a natural group to benchmark the ASVAB because it is a nationally representative sample of the cohort of Americans born 1957–1964. The respondents took the ASVAB in the summer and fall of 1980, following the standard ASVAB procedures. This study of benchmarking the ASVAB using the NLSY–79 is called "Profile of American Youth (PAY–80)."

A major revision of the ASVAB occurred when it shifted from a paper-based test to a computer-based test. The military started to implement the computer-based ASVAB on a large scale in 1996–1997, after about two decades of research and evaluation. The NLSY–97 respondents took the computer-based test, while the NLSY–79 respondents took the paper-based test.

In the paper-based test, all respondents received the same set of questions. In the computer-based test, the next questions that respondents received depend on their answers to previous questions. For example, if a respondent answered a question correctly, then the next question becomes more difficult. This adaptive feature of the computer-based test

[1]For example, in 2015, the military services typically do not accept applicants who score in the bottom 30th percentile in the national AFQT distribution. In addition, DOD requires that at least 60 percent of new enlistees score at the 50th percentile or higher in the national AFQT distribution (Quester and Shuford, 2017).

means that different respondents can receive different sets of questions and with different orderings. The raw count of correct answers is therefore no longer directly comparable across respondents. Instead, item response theory (IRT) models are used to construct estimates of ability and skill (also called "thetas") for each respondent of the computer-based ASVAB. These IRT estimates are supposed to be comparable across respondents.[2]

Due to the test format change, the military needed a new benchmark for the computer-based ASVAB. As the NLSY–97 respondents were 12-17 when first interviewed in 1997, and some were deemed too young for the purpose of benchmarking military enlistees, two other nationally representative samples were identified to complete the computer-based ASVAB during the NLSY–97 screening process. The first sample, the Student Testing Program (STP), consisted of students who expected to be in grade 10–12 in the fall of 1997. Included were many respondents who also participated in the NLSY–97, as well as youth who refused to participate in or were not eligible for the NLSY–97. The second sample, the Enlistment Testing Program (ETP), was a nationally representative sample of youth aged 18–23 as of June 1997. The ASVAB performance of respondents in these two samples (again, which includes some NLSY–97 respondents) was then used to benchmark the computer-based ASVAB for the military.

A.2 Concordance of different formats of ASVAB

A practical issue coming from ASVAB's format change is how to concord the paper-based and computer-based test scores. This is of significant importance for the military because, ideally, the selection criteria into the Armed Forces should be held broadly consistent

[2]Two sections, numerical operations and coding speed, in the computer-based ASVAB are administered in a non-adaptive format (that is, everyone answers the same questions in the same order). The scores of these two sections are therefore not "thetas" estimated from IRT. However, the two sections are still done on computers, so the scores are not directly comparable to the scores of the same sections but under paper format.

before and after the test format change. This is also extremely important for researchers because otherwise the AFQT score and the ASVAB subsection scores, as measures of skills and abilities, are not comparable between the NLSY–79 and the NLSY–97 cohorts (Altonji, Bharadwaj, and Lange, 2012).

Daniel Segall, a researcher at the DOD specializing in psychometrics, developed a mapping between the paper-based and computer-based ASVAB scores (Segall, 1997). He drew a sample of military applicants in two rounds, in 1988 (N=8,040) and from 1990 to 1992 (N=10,379). In each round, one third of the participants were randomly assigned to take the paper-based ASVAB, and the other two–thirds took the computer-based ASVAB. Using the test performance of these military applicants, Segall created a mapping to link each computer-based ASVAB component score to a paper-based ASVAB component score. Since the computer-based ASVAB scores (”thetas” estimated from IRT models) are continuous and the paper-based ASVAB scores (counts of correct answers) are discrete by construction, Segall applied certain smoothing and grouping to the score distributions in the mapping procedure. For further technical details, see Segall (1997).

In their efforts to concord the AFQT score between the NLSY–79 and the NLSY–97, Altonji, Bharadwaj, and Lange (2012) relied heavily on Segall’s mapping. Since the mapping is not publicly available, the authors sent the computer-based ASVAB subsection IRT scores in the NLSY–97 to Segall, who mapped the scores into paper-based scores so that they could directly be compared to the scores of the NLSY–79 respondents. With the scores from Segall in hand, the authors adjusted for one more important difference between the two NLSY cohorts: test-taking ages. The NLSY–79 respondents were around ages 15–23 and the NLSY–97 respondents were around ages 12–18 when they took the ASVAB. On average, ASVAB performance improves as people age, and so it is critical to address the differential test-taking ages both within and cross cohorts.

To construct the mapping across ages, the authors exploited the fact that both cohorts

have a nontrivial share of respondents taking the ASVAB at age 16. Under the (somewhat strong) assumption that a person's *ranking* in the AFQT score distribution does not vary with age, the authors mapped a person at age X (which is not 16) to the score distribution of age 16 by their ranking in the score distribution of age X. For example, if a youth in the NLSY–79 took the test at age 20 and ranked the 25th percentile within the AFQT score distribution of age 20, the youth will be mapped to have the 25th percentile score of the age-16 distribution in the NLSY–79. This relies on the assumption that whoever at the 25th percentile in the score distribution at age 16 will remain at the 25th percentile at age 20. Whether this rank-invariant assumption holds remains to be analyzed and tested. More details can be found in Altonji, Bharadwaj, and Lange (2009).

In addition to converting the computer-based scores in the NLSY–97 to the paper-based scores in the NLSY–79 (as Altonji, Bharadwaj, and Lange (2012) did), another potential approach to compare ASVAB scores between the two formats is to also construct ability estimates (”thetas”) using IRT for the NLSY–79. NLS Program staff created ”thetas” for four sections of ASVAB in the NLSY–79 and made them publicly available.[3] It is not perfectly obvious whether the ”thetas” are directly comparable between the NLSY–79 and the NLSY–97, for at least two reasons. First, it is unclear whether the IRT models and estimation methods used for the two cohorts are the same. Second, even if the models and methods are the same, the raw data taken to the models may still be incomparable due to the very different test formats.

A.3 Different versions of AFQT score

The ASVAB has multiple sections. The AFQT score is a sum of scores from four ASVAB sections. By picking scores from different sections, two versions of the AFQT

[3]The four sections are arithmetic reasoning, math knowledge, paragraph comprehension, and word knowledge.

score have been constructed and used. The AFQT–80, probably the most widely used AFQT score, is the summation of arithmetic reasoning (AR), numerical operations (NO), paragraph comprehension (PC), and word knowledge (WK). The formula is AFQT–80 = AR + 0.5*NO + PC + WK.

In 1989, according to the NLS website, it was realized that the numerical operations section had some design inconsistencies that resulted in unreliable scores (Bureau of Labor Statistics, n.d.(b)). The DOD decided to replace numerical operations with math knowledge (MK) in the construction of the AFQT score. The new score is called AFQT–89. The formula is AFQT–89 = AR + MK + 2*VE. Verbal composite (VE) can be seen as a weighted average of PC and WK with unequal weights. WK receives a higher weight because there are more questions in the WK section.

Different studies have used different versions of the AFQT score. Neal and Johnson (1996) used the AFQT–89 in the published version of their paper, and noted that results are similar using the AFQT–80. Altonji, Bharadwaj, and Lange, 2012 used the AFQT–80 and created the adjusted score that is supposed to be comparable between the NLSY–79 and the NLSY–97. More recent studies have been using their adjusted AFQT–80 score (Castex and Dechter, 2014; Deming, 2017). Although Altonji, Bharadwaj, and Lange, 2012 only did the adjustment for the AFQT–80, their method can be applied to the AFQT–89 and/or the ASVAB subsection scores.

In this dissertation, unless otherwise noted, I used the AFQT–80 score to be consistent with Altonji, Bharadwaj, and Lange (2012), Castex and Dechter (2014), and Deming (2017), who also compare the NLSY–79 with the NLSY–97.

Bibliography

Aaronson, Daniel (1998). "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes". In: *Journal of Human Resources*.

Acemoglu, Daron (2002). "Technical Change, Inequality, and the Labor Market". In: *Journal of Economic Literature*.

Altonji, Joseph, Prashant Bharadwaj, and Fabian Lange (2009). *Constructing AFQT Scores that are Comparable Across the NLSY79 and the NLSY97*.

— (2012). "Changes in the Characteristics of American Youth: Implications for Adult Outcomes". In: *Journal of Labor Economics*.

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

Autor, David (2019). "Work of the Past, Work of the Future". In: *AEA Papers and Proceedings*.

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*.

Autor, David, David Dorn, and Gordon Hanson (2013). "The China Syndrome: Local Labor Market Effects of Import Competition in the United States". In: *American Economic Review*.

Barsky, Robert, John Bound, Kerwin Charles, and Joseph Lupton (2002). "Accounting for the Black-White Wealth Gap: A Nonparametric Approach". In: *Journal of the American Statistical Association*.

Baumrind, Diana (1991). “The Influence of Parenting Style on Adolescent Competence and Substance Use”. In: *The Journal of Early Adolescence*.

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*.

Beaudry, Paul, David Green, and Benjamin Sand (2016). “The Great Reversal in the Demand for Skill and Cognitive Tasks”. In: *Journal of Labor Economics*.

Becker, Gary (1964). *Human capital: A theoretical and empirical analysis with special reference to education*. University of Chicago press.

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*.

Bureau of Labor Statistics (1992). *NLSY79 Profiles of American Youth: Addendum to Attachment 106*. Tech. rep.

— (2019). *Employment Projections*. `https://www.bls.gov/emp/tables/civilian-labor-force-summary.htm`. Accessed: 2019-10-01.

— (n.d.[a]). *Administration of the CAT-ASVAB (NLSY97)*. `https://www.nlsinfo.org/content/cohorts/nlsy97/topical-guide/education/administration-cat-asvab-0`. Accessed: 2020-06-23.

— (n.d.[b]). *Aptitude, Achievement Intelligence Scores (NLSY79)*. `https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/education/aptitude-achievement-intelligence-scores`. Accessed: 2020-06-23.

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

Charles, Camille (2003). “The Dynamics of Racial Residential Segregation”. In: *Annual Review of Sociology*.

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, and Sonya Porter (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*.

Coulson, Edward and Herman Li (2013). “Measuring the external benefits of homeownership”. In: *Journal of Urban Economics*.

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

Cunha, Flavio and James Heckman (2007). “The Technology of Skill Formation”. In: *American Economic Review*.

Cunha, Flavio, James Heckman, Lance Lochner, and Dimitriy Masterov (2006). “Interpreting the Evidence on Life Cycle Skill Formation”. In: *Handbook of Economics of Education*. Elsevier.

Cutler, David and Edward Glaeser (1997). "Are Ghettos Good or Bad?" In: *The Quarterly Journal of Economics*.

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*.

Darity, William and Patrick Mason (1998). "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender". In: *Journal of Economic Perspectives*.

Defense Manpower Data Center (2006). *ASVAB Technical Bulletin*. Tech. rep.

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

Dillon, Eleanor and Jeffrey Smith (2020). "The Consequences of Academic Match between Students and Colleges". In: *Journal of Human Resources*.

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

DiPasquale, Denise and Edward Glaeser (1999). "Incentives and Social Capital: Are Homeowners Better Citizens?" In: *Journal of Urban Economics*.

Doepke, Matthias, Giuseppe Sorrenti, and Fabrizio Zilibotti (2019). "The Economics of Parenting". In: *Annual Review of Economics*.

Doepke, Matthias and Fabrizio Zilibotti (2017). "Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission". 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*.

Economic Policy Institute (2020). "Schools are still segregated, and black children are paying a price". In: URL: `https://www.epi.org/publication/schools-are-still-segregated-and-black-children-are-paying-a-price/`.

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, James Heckman, Duncan Ermini Leaf, and María José Prados (2020). "Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program". In: *Journal of Political Economy*.

García, Jorge Luis, James Heckman, and Anna Ziff (2019). "Early childhood education and crime". In: *Infant Mental Health Journal*.

Griliches, Zvi (1977). "Estimating the Returns to Schooling: Some Econometric Problems". In: *Econometrica*.

— (1986). "Economic data issues". In: *Handbook of Econometrics*. Elsevier.

Griliches, Zvi and William Mason (1972). "Education, Income, and Ability". In: *Journal of Political Economy*.

Grogger, Jeff (1992). "Arrests, Persistent Youth Joblessness, and Black/White Employment Differentials". In: *Review of Economics and Statistics*.

— (1998). "Market Wages and Youth Crime". In: *Journal of Labor Economics*.

Heckman, James (1981). "Heterogeneity and state dependence". In: *Studies in Labor Markets*.

— (1998). "Detecting Discrimination". In: *Journal of Economic Perspectives*.

Heckman, James (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, Thomas Lyons, and Petra Todd (2000). "Understanding Black-White Wage Differentials, 1960-1990". In: *American Economic Review*.

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*.

Ing, Pamela, Carole Lunney, and Randall Olsen (2012). "Reanalysis of the 1980 AFQT Data from the NLSY79". In: *Center for Human Resource Research Working Paper*.

Jacob, Brian and Jesse Rothstein (2016). "The Measurement of Student Ability in Modern Assessment Systems". In: *Journal of Economic Perspectives*.

Johnson, Janna and Sam Schulhofer-Wohl (2019). "Changing Patterns of Geographic Mobility and the Labor Market for Young Adults". In: *Journal of Labor Economics*.

Johnson, William, Yuichi Kitamura, and Derek Neal (2000). "Evaluating a Simple Method for Estimating Black-White Gaps in Median Wages". In: *American Economic Review*.

Junker, Brian, Lynne Schofield, and Lowell Taylor (2012). “The use of cognitive ability measures as explanatory variables in regression analysis”. In: *IZA Journal of Labor Economics*.

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

Kahn, Lisa and David Deming (2017). “Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals”. In: *NBER Working Paper*.

Kain, John (1968). “Housing Segregation, Negro Employment, and Metropolitan Decentralization*”. In: *The Quarterly Journal of Economics*.

Kondo, Ayako (2015). “Differential effects of graduating during a recession across gender and race”. In: *IZA Journal of Labor Economics*.

Lang, Kevin and Jee-Yeon K. Lehmann (2012). “Racial Discrimination in the Labor Market: Theory and Empirics”. In: *Journal of Economic Literature*.

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

Lareau, Annette (1987). “Social Class Differences in Family-School Relationships: The Importance of Cultural Capital”. In: *Sociology of Education*.

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

Light, Audrey and Manuelita Ureta (1995). “Early-Career Work Experience and Gender Wage Differentials”. In: *Journal of Labor Economics*.

Maccoby, E. and J. Martin (1983). “Socialization in the Context of the Family: Parent-Child Interaction”. In: *Handbook of Child Psychology*.

Massey, Douglas and Nancy Denton (1993). *American apartheid: Segregation and the making of the underclass*. Harvard University Press.

McAdoo, Harriette (2002). “African American parenting”. In: *Handbook of Parenting*.

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

Mincer, Jacob (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research.

Neal, Derek (2004). “The Measured Black-White Wage Gap among Women Is Too Small”. In: *Journal of Political Economy*.

— (2006). “Why Has Black–White Skill Convergence Stopped?” In: *Handbook of the Economics of Education*. Elsevier.

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:

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

New York Times (2018). “Extensive Data Shows Punishing Reach of Racism for Black Boys”. In: URL: `https://www.nytimes.com/interactive/2018/03/19/upshot/race-class-white-and-black-men.html`.

— (2019). “Still Separate, Still Unequal: Teaching about School Segregation and Educational Inequality”. In: URL: `https://www.nytimes.com/2019/05/02/learning/lesson-plans/still-separate-still-unequal-teaching-about-school-segregation-and-educational-inequality.html`.

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

Ondrich, Jan, Stephen Ross, and John Yinger (2000). “How Common is Housing Discrimination? Improving on Traditional Measures”. In: *Journal of Urban Economics*.

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

Pew Research Center (2017). “It’s becoming more common for young adults to live at home and for longer stretches”. In: URL: https://www.pewresearch.org/fact-tank/2017/05/05/its-becoming-more-common-for-young-adults-to-live-at-home-and-for-longer-stretches/.

— (2018). “Millennials are largest generation in the U.S. labor force”. In: URL: https://www.pewresearch.org/fact-tank/2018/04/11/millennials-largest-generation-us-labor-force/.

— (2019a). “Defining generations: Where Millennials end and Generation Z begins”. In: URL: https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/.

— (2019b). “Gen Z, Millennials and Gen X outvoted older generations in 2018 midterms”. In: URL: https://www.pewresearch.org/fact-tank/2019/05/29/gen-z-millennials-and-gen-x-outvoted-older-generations-in-2018-midterms/.

— (2020). “As Millennials Near 40, They’re Approaching Family Life Differently Than Previous Generations”. In: URL: https://www.pewsocialtrends.org/2020/05/27/as-millennials-near-40-theyre-approaching-family-life-differently-than-previous-generations/.

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*.

Quester, Aline and Robert Shuford (2017). *Population Representation in the Military Services: Fiscal Year 2015 Summary Report*. Tech. rep.

Reardon, Sean and Ann Owens (2014). “60 Years After Brown: Trends and Consequences of School Segregation”. In: *Annual Review of Sociology*.

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*.

Rothstein, Jesse (2019). “The Lost Generation? Scarring after the Great Recession”. In: *Working Paper*.

Sands, William, Brian Waters, and James McBride, eds. (1997). *Computerized Adaptive Testing: From Inquiry to Operation*. American Psychological Association.

Schofield, Lynne (2014). “Measurement error in the AFQT in the NLSY79”. In: *Economics Letters*.

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*.

Segall, Daniel (1997). “Equating the CAT-ASVAB”. In: *Computerized Adaptive Testing: From Inquiry to Operation*. American Psychological Association.

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/`.

— (2018). “School Segregation Is Not a Myth”. In: URL: `https://www.theatlantic.com/education/archive/2018/03/school-segregation-is-not-a-myth/555614/`.

Thompson, Owen (2018). “The Determinants of Racial Differences in Parenting Practices”. In: *Journal of Political Economy*.

Tobin, James and William Brainard (1976). “Asset markets and the cost of capital”. In: *Cowles Foundation Discussion Papers*.

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*.

Uzawa, Hirofumi (1969). “Time preference and the Penrose effect in a two-class model of economic growth”. In: *Journal of Political Economy*.

Washington Post (2018). “Same family income, same street, but the black boy still grows up to earn less”. In: URL: `https://www.washingtonpost.com/news/wonk/wp/2018/03/19/same-family-income-same-street-but-the-black-boy-still-grows-up-to-earn-less/`.

Weinberger, Catherine (2014). “The Increasing Complementarity between Cognitive and Social Skills”. In: *Review of Economics and Statistics*.

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.

Yagan, Danny (2019). “Employment Hysteresis from the Great Recession”. In: *Journal of Political Economy*.

Yinger, John (1995). *Closed doors, opportunities lost: The continuing costs of housing discrimination*. Russell Sage Foundation.

Yitzhaki, Shlomo (1996). “On Using Linear Regressions in Welfare Economics”. In: *Journal of Business & Economic Statistics*.

Zhao, Bo, Jan Ondrich, and John Yinger (2006). “Why do real estate brokers continue to discriminate? Evidence from the 2000 Housing Discrimination Study”. In: *Journal of Urban Economics*.