DYNAMIC SKILL DEVELOPMENT AND LABOR MARKET OUTCOMES

By

John Brauer

A dissertation submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

(Economics)

at the

UNIVERSITY OF WISCONSIN – MADISON

2019

Date of final oral examination: 6/10/2019

The dissertation is approved by the following members of the Final Oral Committee:

John F. Kennan, Professor, Economics

Christopher Taber, Professor, Economics

Chao Fu, Professor, Economics

Robert Meyer, Professor, Public Affairs

Abstract

The first chapter investigates the presence of statistical discrimination in the labor market. The Children of the NLSY79 data are used to link early-age home environment measures to educational attainment measures and labor market outcomes. While both black and white children with higher measured home inputs sort into higher levels of educational attainment, this positive sorting pattern is significantly stronger for black children. Estimates also reveal that, after controlling for a variety of skill measures, the residual black-white wage gap is large for high school dropouts and narrows rapidly with additional educational attainment. For college-goers, measured skills can account for the entire black-white wage gap. These patterns are consistent with a scenario in which employers use both race and education credentials to form expectations about elements of worker productivity formed through early-age inputs. Under plausible and partially testable identifying assumptions, the results imply that a portion of the black-white wage gap for low-education workers reflects statistical discrimination in the labor market.

Skill development in college and on the job can depend not only on the quality of investments but also on the order in which these investments are made. The second chapter explores which types of occupational investments complement college best when performed before college entry and which types are more productive after college completion. A learning-by-doing model with both college entry timing and early-career occupation choices produces several key insights. Data from the NLSY79 are linked with abstract and routine occupational task content data, and relationships between college entry timing, early-career occupation choices, and future earnings trajectories are documented. Estimates suggest that abstract-intensive occupations are more beneficial for skill development just after college, whereas routine-intensive occupations are more beneficial for skill development before college. Accordingly, delayed college entrants choose more routine-intensive early-career occupations, and immediate college entrants choose more abstract-intensive early-career occupations. The results also indicate that high school graduates with high levels of abstract skills face the largest penalty for delaying college entry.

Acknowledgements

I would like to thank John Kennan for providing me with direction and unwavering support throughout my time as a graduate student. Discussions about big ideas, small details, and everything in between have been truly illuminating and will continue to impact the work that I do for the rest of my career. I also wish to thank Chris Taber for being so generous with his time and always refreshing me with his excitement, perspective, and insight. I am indebted to my other committee members, Chao Fu and Rob Meyer, for their support and sound advice. I also appreciate the many other members of the economics department faculty who have helped me along the way.

This would not have been possible without my family. I would not be here today without my incredible parents, Bill and Sylvia. Your unending support throughout this journey has been essential to my success. My sisters, Laura and Andrea, have also been an enormous source of joy and solace. I have been extremely fortunate to have had my family so close by for all these years. Auntie Mary Alice, Uncle John, Cousin Mike, and others also provided important words of wisdom.

Adam, Pedro, Moheeb, Gabriel, Daniel, and other friends made in the program, thanks for your warm friendship and inspiring scholarship. Finally, thanks to Alex, Daniel, Ben, Michael, Zach, Adele, and all of my wonderful friends for the fun and laughter.

List of Figures

List of Tables

Contents

Chapter 1

Labor Market Discrimination: Race and Education as Signals of the Early-Age Skill Environment

1.1 Introduction

Black-white wage gaps in the United States are well-documented and widely studied. While these gaps have persisted over many generations, the importance of discrimination in producing racial wage gaps and racial disparities in other labor market outcomes has been widely debated among social scientists. For one, racial disparities in measured skill bundles can account for a significant portion of racial wage gaps. Neal and Johnson (1996) demonstrate that controlling for the Armed Forces Qualification Test score alone accounts for a large portion of the black-white wage gap for men in the NLSY79 data and virtually all of the gap for women.

If there exist racial disparities in important skills and other wage determinants that are unobserved by the econometrician, then such estimates of racial wage gaps may misstate the extent of racial discrimination operating in the labor market. Researchers have attempted to overcome this fundamental identification problem in a variety of ways, including development of structural models of the labor market and quasi-experimental audit studies.[1] Despite these efforts, there is not a clear consensus on the extent to which discrimination contributes to black-white wage gaps.

This chapter provides several contributions to the literature on labor market discrimination. First, this chapter takes advantage of a unique set of data that links measures of early-age home environment to later test scores, educational attainment measures, and early-career labor market outcomes. The Children of the NLSY79 (CNLSY79) contains multiple cohorts of children with valid and comprehensive measures of early-age home inputs from birth in addition to later-age math and reading test scores, high school and postsecondary education attainment, and early-career hourly wages. Todd and Wolpin (2003) demonstrate that these home environment measures serve as inputs to crucial stages of cognitive development. Todd and Wolpin (2007), Cunha et al (2010), and others have demonstrated that these early-age environments are an important input for the development of social and emotional skills. A large body of work has also linked cognitive and noncognitive skill measures to labor market outcomes. To my knowledge, no other chapter has directly investigated the link between early-age home inputs, educational attainment, and labor market outcomes. Rather, the existing literature typically infers the effects of early-age home inputs on labor market outcomes indirectly through their effect on skill measures.

It has been widely documented that measures of home inputs to skill production are significantly lower, on average, in black households.[2] Kan and Tsai (2005) document that

[1]See Bowlus and Eckstein (2002) and Bertrand and Mullanaithan (2004), respectively, for some prominent examples.

[2]Todd and Wolpin (2007) and Thompson (2018) are just a few examples.

more stimulating home environments are generally associated with higher educational attainment. This chapter documents significant variation in racial gaps in measured home inputs *across* eventual educational attainment levels. Using a comprehensive measure of the early-age home environment, the average black level of measured home inputs lags the average white level of the same measure at every level of educational attainment. At the lowest levels of educational attainment, the racial gap in home inputs is particularly large. Black high school dropouts with no GED have, on average, almost a full standard deviation lower levels of home inputs than their white counterparts. This racial gap in home inputs narrows with educational attainment as black workers received markedly higher levels of home inputs. For college-goers, the racial gap in home inputs is less than half of a standard deviation.

The early-age home environment serves as a crucial input for the development of a wide set of skills, some of which may not be fully observed by employers forming expectations about the productivity of a worker. In the presence of statistical discrimination, employers can use easy-to-observe correlates, such as educational attainment credentials and racial identity, to form beliefs about hard-to-observe skills developed in early ages. In this view, employers form particularly negative beliefs about the productivity levels of black high school dropouts relative to white high school dropouts. This racial difference in beliefs narrows with increased levels of educational attainment, since the sorting patterns of home inputs by educational attainment levels are stronger for black workers than for white workers. The value of educational attainment as a signal of these hard-to-observe skills will be higher for black workers, and the labor market returns to educational attainment will therefore be greater for black workers.

This chapter documents a significant difference in the measured wage returns to education for black and white workers up to 13 years of educational attainment. Controlling for home input measures as well as several other measures of skill, the residual racial wage gap is largest at the lowest levels of education. This residual wage gap decreases rapidly for workers who attain higher levels of education, and measured skills explain the entire racial wage gap for college-goers. These patterns, along with the documented racial difference in educational sorting by home inputs, are consistent with the simple view that employers statistically discriminate on the basis of both race and educational attainment. Assumptions about important skills that are unobserved by the econometrician are explored in the context of the model, and several robustness checks are performed. Results suggest that a significant portion of the racial wage gap for low education workers is the result of employer statistical discrimination.

The chapter proceeds as follows. Section 2 describes the novelty of this approach in relation to the existing literature. Section 3 presents some basic patterns in the data that motivate the model. Section 4 presents a model of statistical discrimination that directly addresses the role of productivity elements unobserved to the econometrician. Section 5 provides an overview of the data used in the empirical analysis. Section 6 presents the main results and Section 7 discusses some alternative mechanisms. Section 8 discusses applications of these results to policy and Section 9 concludes.

1.2 Relationship to the Existing Literature

Arrow's (1973) seminal contribution to the discrimination literature lays a conceptual foundation of statistical discrimination. Arrow (1973) makes the crucial distinction between investments in human capital that are easy to measure and investments that are "more subtle types of personal deprivation and deferment of gratification which lead to the habit and action of thought that favor good performance in skilled jobs, steadiness, punctuality, responsiveness, and initiative." If employers form relatively low expectations about hard-to-observe skills possessed by black workers, then black workers will be rewarded less for investment in such skills. In turn, black workers invest less in these skills, confirming the low expectations formed by employers. A large literature has used this basic framework in an attempt to better understand the source of black-white wage gaps.[3] In this literature, a common theme arises in which equilibrium black-white wage gaps due to statistical discrimination persist in tandem with lower returns to skill investments for black workers.

Neal (2006) points out that data from recent decades generally does not support the hypothesis that black workers face lower returns to skill investments. This point has been interpreted to suggest that the standard view of statistical discrimination is not as empirically relevant as it was in the time of its conception. This chapter provides an alternative view of statistical discrimination that does not necessarily predict lower returns to hard-to-observe skills for black workers. If firms use race and educational attainment as signals of hard-to-observe skills, then the signaling value of education will be higher for black workers precisely because the sorting patterns into education on hard-to-observe skills are stronger for black workers. These sorting patterns are consistent with patterns observed in the data on early-age home inputs presented in this chapter. Crucially, this view of statistical discrimination does not rely on lower returns

[3]Aiger and Cain (1977), Lundberg and Startz (1983), and Coate and Loury (1993) are just a few prominent examples.

skill investments for black workers. The view also predicts that the labor market returns to educational attainment are actually higher for black workers, which is generally supported by data from more recent decades.

Lang and Manove (2011) present a related model in which racial differences in the signaling value of education arise in equilibrium. In order to generate this prediction, their model imposes particular race and education-specific signaling technologies. Firm signals of ability are especially noisy for low-education black workers, and black workers can reveal true ability much more clearly with increased educational attainment. By construction, racial differences in the signal noise imply that the returns to unobserved skills are lower for black workers, a prediction that is generally not supported by the data.

In another empirical application of statistical discrimination, Altonji and Pierret (2001) consider the behavior of employers as they learn about true worker productivity over time. When the worker enters the labor force, employers use easy-to-observe correlates of productivity (such as race or education) to form expectations of productivity. With worker experience, employers learn about true worker productivity and rely less on the productivity signal in wage assignment. Crucial to identification of the model is a measure of skill that is observed by the econometrician but not observed by firms, and Altonji and Pierret (2001) primarily use AFQT to measure such skills in their analysis. However, this approach will not accurately capture employer learning if firms can observe or evaluate most of the important skills associated with AFQT at or near the time of labor market entry. This chapter takes a new approach, instead using measures of the early-age home environment which are more likely to feed into important skills that are difficult for employers to assess at the time of labor market entry.

Arcidiacono et al (2010) present a simple but novel extension of the Altonji and Pierret (2001) framework. Using the same NLSY79 data, they separately consider a sample of high school graduates who never enrolled in college and a sample of college graduates. Patterns in the data suggest that productivity of college graduates (as measured by AFQT) is revealed very quickly to employers. Further, the racial wage gap is eliminated and is actually reversed for college graduates after controlling for AFQT. In contrast, productivity is revealed more gradually with labor market experience for high school graduates. The racial wage gap for high school graduates starts out large and significant and actually increases somewhat with labor market experience. Put differently, employers learn about elements of productivity measured by AFQT over time but this information revelation does not result in a narrowing of the racial wage gap for high school graduates. This fact is inconsistent with the presence of statistical discrimination on the basis of race as predicted by the basic model of Altonji and Pierret (2001).

In the context of the results in Arcidiacono et al (2010), this chapter suggests that the hard-to-observe elements of productivity which cause firms to assign wages differentially on the basis of race may not be properly captured by the AFQT score. Instead, this chapter finds that variation in those skills may be more accurately captured by variation in crucial early-age investments that provide a foundation for the development of a wide array of skills. In this sense, this chapter complements the existing literature by providing a new interpretation of the residual black-white wage gap for workers with low levels of educational attainment.

1.3 Motivating Data Patterns

In order to motivate the basic mechanism, some data patterns are first presented. The sample used in this chapter includes both white and black non-Hispanic males in the Children of the NLSY79 sample. This section considers relationships between the following measures for each child in the sample: a race indicator, a HOME score constructed from the Home Observation Measurement of the Environment items between ages 0 and 9, Peabody Individual Achievement Test (PIAT) standardized math and reading scores, educational attainment measures, and log hourly wages recorded between ages 22 and 27.[4]

The HOME score is constructed from several dichotomous variables that measure the adequacy (or inadequacy) of several aspects of the child's home environment. These items cover a relatively wide set of important dimensions of the home environment, ranging from measures of cognitive stimulation to emotional support and types of discipline for misbehavior.[5] When necessary, these items have open-ended responses and are mapped into a binary outcome measure in order to allow for some flexibility in the types of behaviors associated with a stimulating home environment.

Table 1 presents a few summary statistics for racial differences in the standardized HOME score, the PIAT math and reading scores, and various educational attainment measures. Average levels of all these measures are lower for blacks in the sample. The overall racial gap in HOME score distributions is just over one half of a standard deviation. The racial gaps in math and reading PIAT test scores are both large. The racial

[4] More details regarding the sample and variable construction are discussed in Section 5.

[5] For a more detailed discussion of the construction, reliability, and use of the HOME score, see Caldwell and Bradley (1984).

gap in educational attainment is also significant. Blacks attain about one year less education on average than whites, are 12% less likely to obtain a high school diploma, and are 18% less likely to enroll in college.

Table 1: Average Skill Measures by Race

Skill Measure	White	Black	Difference
HOME Score (standard units)	.13	-.45	.58
Math PIAT (standard units)	.35	-.48	.83
Reading PIAT (standard units)	.22	-.49	.71
Highest Grade Completed	13.4	12.5	0.9
High School Diploma	87%	75%	12%
Any College	64%	46%	18%

Thompson (2018) also documents sizable racial differences in certain objective HOME items, including whether the mother reads to the child 3+ times per week, whether the home has 10+ children's books, and whether the mother responded to the child's speech during the interview. Other researchers have documented similarly large racial gaps in the home environment using alternative measures.[6]

Determinants of racial differences in educational attainment have been widely studied. These racial differences can be attributed to significant differences in long-run family background measures and measures of skills near the time of labor market entry.[7] In contrast, possible determinants of racial differences in the child's home environment have been subject to fewer studies.[8] Table 2 reveals that maternal AFQT and permanent household income can together account for 75% of the gap in measured HOME score. After including an indicator for the mother's marital status, the remaining gap is

[6]For specific examples, see Brooks-Gunn and Markman (2005) and Ferguson (2005).

[7]Carneiro et al (2005) and Lang and Manove (2011) provide compelling evidence.

[8]Thompson (2018) documents these gaps and provides one possible explanation for variation in HOME scores over time and across race and region in the US.

statistically insignificant at the 90% level.

Table 2: Black-White Gaps in HOME Score, Ages 0-9

	(1)	(2)	(3)	(4)	(5)
Black	-0.386	-0.162	-0.203	-0.106	-0.0728
	(0.0341)	(0.0378)	(0.0341)	(0.0367)	(0.0377)
Maternal AFQT		0.222		0.127	0.127
		(0.0169)		(0.0174)	(0.0174)
Household Income/10,000			0.0898	0.0747	0.0715
			(0.00513)	(0.00546)	(0.00546)
Marital Status					0.111
					(0.0310)
Observations	2732	2670	2719	2659	2659
R-squared	0.415	0.450	0.472	0.483	0.489

Standard errors in parentheses. Dependent variable is the standardized mean of all HOME scores recorded between ages 0 and 9. All specifications also control dummies for southern regional status, birth order, maternal age at birth, ages HOME score observed, and total number of HOME observations.

In this light, the racial gap in HOME scores can be seen largely as a reflection of racial inequality in outcomes of the previous generation that translate directly into crucial differences in the child's home environment. In addition to differences in the time and money inputs that shape the home environment for children, poverty can cause a significant amount of parental stress that also affects parenting practices.[9] As such, it is unlikely that families living in poverty could substantially improve the skill inputs that they provide to their children in the absence of intervention.

Figure 1 explores racial differences in average HOME scores for children who ultimately received different educational credentials. There are a few notable features here.

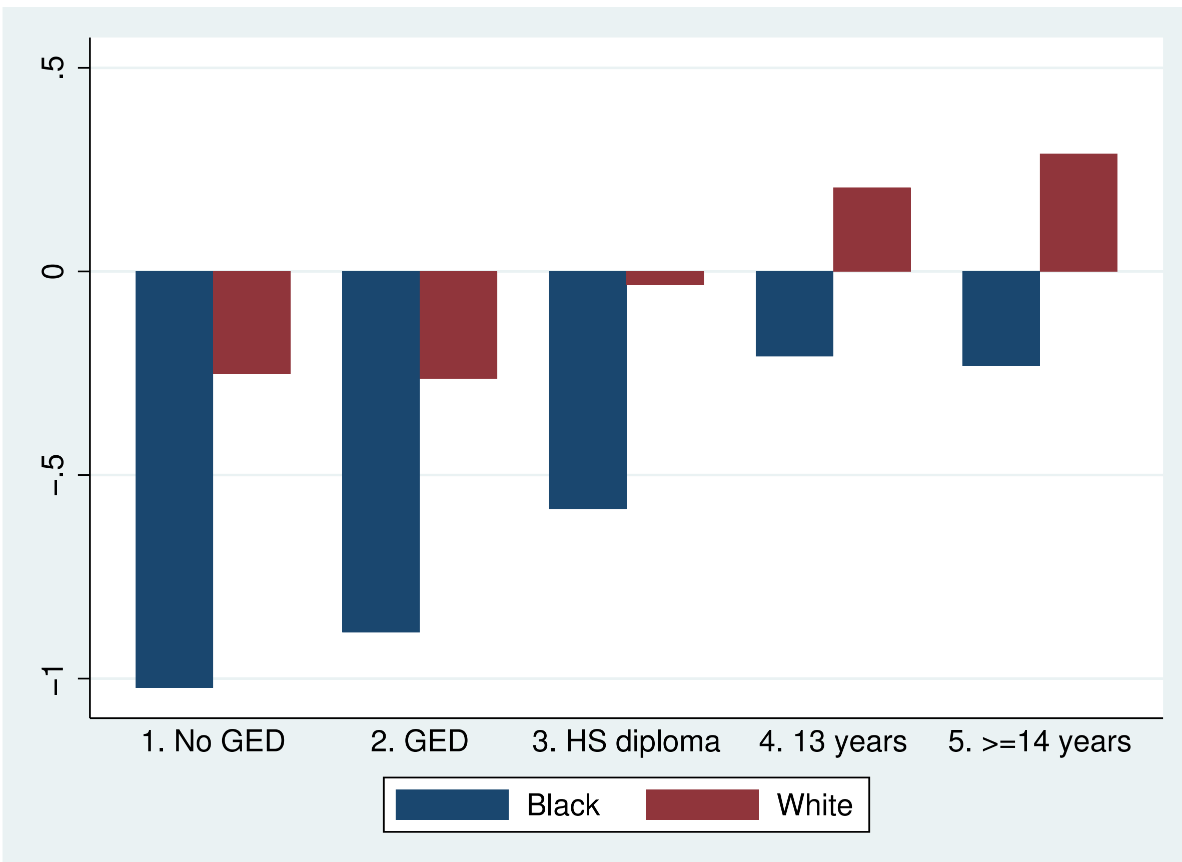
[9]The separate effects of parental stress resulting from long-term poverty is detailed in a literature surrounding the Family Stress Model developed in Conger et al (2000).

The data exhibit a pattern of positive HOME score sorting into educational outcomes for both black and white workers, but this positive sorting is stronger for black workers than white workers. Average racial differences in the HOME score are especially striking at the lowest levels of educational attainment. The average HOME score for black high school dropouts are significantly lower than for any other group. Black workers who received a high school diploma but did not attend any college had substantially higher HOME scores than black workers who did not receive a high school diploma. However, these levels are still significantly lower on average than any group of whites and all groups of blacks with higher levels of education. Black workers who attended college also had a substantially higher average HOME score, while white workers who attended college saw relatively modest increases. Finally, average HOME score for blacks who attended at least some college are more comparable to the average HOME scores of white workers.[10]

In order to further examine race-specific correlations between HOME score and educational attainment, Table 3 displays estimates from a regression of HOME scores on race, educational attainment, and an interaction between race and education. Column (1) uses highest grade completed as the measure of educational attainment and Column (2) uses the same educational categories displayed in Figure 1. The estimates from Table 3 reveal that the HOME score is positively correlated with educational attainment and negatively correlated with race. The estimates of coefficients on the interaction terms provide evidence of racial differences in the education sorting patterns. Positive and statistically significant estimates of these interaction coefficients provides more evidence that positive HOME score sorting into education is stronger for black workers than for

[10]Figure 4 in Appendix B displays average HOME scores by race and highest grade completed. The patterns there are largely consistent with those in Figure 1. In particular, the racial gap in HOME scores narrows with increases in years of schooling completed from 9 to 13 years of schooling.

Figure 1: Average HOME Scores by Race and Education



white workers.

Table 3: Race-Specific Correlations Between HOME and Education

	(1)	(2)
Black	-0.576	-0.566
	(0.0482)	(0.0458)
Ed Category	0.136	
	(0.0191)	
Black*Ed Category	0.0717	
	(0.0309)	
Highest Grade		0.0905
		(0.0103)
Black*Highest Grade		0.0458
		(0.0209)
Constant	0.0330	0.0418
	(0.0312)	(0.0281)
Observations	1701	1524
R-squared	0.2472	0.2503

Standard errors in parentheses. Dependent variable is the HOME score.

The HOME score is fundamentally an ordinal measure, and the narrowing racial gaps in home inputs are not preserved if certain monotonic transformations of the HOME score are considered in its place. However, observed hourly wages in the data are strongly log-linear in the HOME score across a wide interval of HOME scores, and the slope of this log-linear relationship does not differ by racial group. Motivated by these patterns, the model presented below is specified with home inputs entering the true productivity process as a log-linear component. This specification has a clear interpretation: HOME score has a constant skill price in the labor market.

It should also be noted that this general pattern of racial differences in educational sorting present in the HOME score is not present when considering other measures of skill. Figure 5 in Appendix B displays racial gaps in PIAT math and reading scores across educational attainment levels. Unlike the patterns of average HOME scores documented in Figure 1, racial gaps in PIAT math scores remain largely constant around three-quarters of a standard deviation across educational attainment levels. For both black and white workers, college attendance is associated with much higher math PIAT scores. Racial gaps in the PIAT reading score narrow somewhat with educational attainment up to receipt of the high school diploma. However, college attendance is associated with much higher reading PIAT scores for both black and white workers.[11]

If employers cannot perfectly observe elements of productivity that depend crucially on the early-age home environment, they may use easily observable correlates such as race and education credentials to form expectations about these skills. The value of attaining important educational milestones (as a signal of these home inputs) will therefore be larger for black workers up to 13 years of educational attainment. In the next section, a simple model is developed that provides a method that allows for a more closely examination of these basic predictions.

1.4 A Simple Model of Statistical Discrimination

Assume that true worker productivity $p^\star$ can be decomposed in the following way.

$$log(p^\star) = a + g(s) + \eta + \epsilon$$

[11]Arcidiacono et al (2010) also demonstrate that racial gaps in AFQT scores do not vary substantially across educational attainment levels.

Here, a represents measured home inputs, s represents measured educational attainment, η is represents variation in productivity inputs from outside the home (such as school and neighborhood quality) that are observed by employers, and ϵ is a portion of productivity not observed by employers or the econometrician. The unobserved productivity shocks are lognormally distributed. That is, it is assumed that $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$.

Employers observe a worker's race r, educational attainment s, the contribution η, and a noisy signal $\tilde{q}$ of the unobserved portion of productivity $a + \epsilon$ given by

$$\tilde{q} = a + \epsilon + u$$

where $u \sim \mathcal{N}(0, \sigma_u^2)$. All employers observe the same signal for a given worker, and workers observe exactly what employers observe, both standard assumptions in the statistical discrimination literature and in the employer learning literature.[12]

It should be noted that the distribution of η may depend upon r and s but is assumed to be independent of a. If η and a were correlated, then η would provide some information about a. Instead, this information revelation is modeled as a signal $\tilde{q}$, distinct from η. This modeling choice is relatively innocuous and allows for a clearer interpretation of each distinguishable component of productivity.

Define $\bar{a}_r(s) = E[a|r, s]$, the employers' expectations of the unobserved portion of log productivity of a worker belonging to racial group r and having educational attainment s prior to the revelation of the productivity signal $\tilde{q}$. Note that employers' prior beliefs about this unobserved portion for a given worker equals the average home inputs in the race-education group of the worker. Thus, sorting patterns of a into s for different r

[12]Analogous assumptions are adopted by Arrow (1973), Lang and Manove (2011), and Altonji and Pierret (2001).

crucially impact these beliefs.

$I_1 = \{r, s, \eta, \tilde{q}, \bar{a}_r(s)\}$ denotes the employers' common information set, and $I_2 = \{r, s, a, \bar{a}_r(s)\}$ denotes the econometrician's information set. To employers, the expectation of $log(p^\star)$ is a weighted average of the signal and the prior of the unobserved portion, plus the observed portion $g(s) + \eta$.

$$E[log(p^\star)|I_1] = \lambda\tilde{q} + (1-\lambda)\bar{a}_r(s) + g(s) + \eta$$

where

$$\lambda = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_u^2}$$

The conditional variance of $log(p^\star)$ is given by

$$\sigma^2[log(p^\star)|I_1] = (1-\lambda)\sigma_\epsilon^2$$

By properties of the lognormal distribution, it follows that

$$log(E[p^\star|I_1]) = E[log(p^\star)|I_1] + \frac{1}{2}\sigma^2(log(p^\star)|I_1)$$

If worker output is not contractible and the market is perfectly competitive, then all employers offer a wage equal to expected worker productivity.

$$log(w) = log(E[p^\star|I_1]) = \lambda\tilde{q} + (1-\lambda)(\bar{a}_r(s) + \frac{1}{2}\sigma_\epsilon^2) + g(s) + \eta$$

Forming an expectation conditional on the econometrician's information set I_2 gives the following expression.

$$E[log(w)|I_2] = \lambda a + (1-\lambda)(\bar{a}_r(s) + \frac{1}{2}\sigma_\epsilon^2) + g(s) + E[\eta|s,r]$$

The expectation of η does not condition on a since above, log productivity is deconstructed so that η and a are independent.

There are two contributions to the difference in conditional log wages, which can be estimated using a standard ordinary-least squares residual racial log wage gap.

$$E[log(w)|a,s,W] - E[log(w)|a,s,B] = \underbrace{(1-\lambda)(\bar{a}_W(s) - \bar{a}_B(s))}_{\text{statistical discrimination}} + \underbrace{(E[\eta|s,W] - E[\eta|s,B])}_{\text{non-home inputs}}$$

In this context, the standard identification problem for interpreting ordinary least-squares estimates of racial wage gaps arises. The first term is due to statistical discrimination, while the second term is due to average racial differences in all other productivity elements observed by employers. If η is correlated with race at some education level, then the two contributions on the racial wage gap cannot be separately identified unless the econometrician can observe all contributions to η.

Rather than focusing on estimates of the measured racial wage gap as much of the previous literature on racial wage discrimination has done, it is informative to instead turn to racial differences in the measured wage returns to educational attainment. In the presence of statistical discrimination, this signaling value of education can differ by race when employers use both race and education as signals. This can drive a wedge between the wage returns to education for black and white workers. Again, the difference in the measured wage returns to education may in part reflect productivity differences across race and education not captured by home inputs. The value of this model is that it provides precise conditions on these non-home productivity inputs that, when satisfied, imply that the difference in returns to education are a result of statistical discrimination.

To see this more clearly, note that the racial difference in the measured returns to education can be expressed as

$$\frac{\partial E[log(w)|a,s,B]}{\partial s} - \frac{\partial E[log(w)|a,s,W]}{\partial s} = \underbrace{(1-\lambda)(\bar{a}'_B(s) - \bar{a}'_W(s))}_{\text{statistical discrimination}} + \underbrace{(\frac{\partial E[\eta|s,B]}{\partial s} - \frac{\partial E[\eta|s,W])}{\partial s})}_{\text{non-home inputs}}$$

Making the critical assumption $\frac{\partial E[\eta|s,W]}{\partial s} \geq \frac{\partial E[\eta|s,B]}{\partial s}$ implies that

$$\frac{\partial E[log(w)|a,s,B]}{\partial s} - \frac{\partial E[log(w)|a,s,W]}{\partial s} \geq (1-\lambda)(\bar{a}'_B(s) - \bar{a}'_W(s))$$

It is also clear that $1-\lambda > 0$ if and only if employers cannot perfectly observe a. The main proposition follows directly.

Proposition: *Let $S = [s_L, s_H]$ be a range of education levels over which*

1. $\bar{a}'_B(s) > \bar{a}'_W(s)$

2. $\frac{\partial E[\eta|s,W]}{\partial s} \geq \frac{\partial E[\eta|s,B]}{\partial s}$

Then, if $\frac{\partial E[log(w)|a,s,B]}{\partial s} > \frac{\partial E[log(w)|a,s,W]}{\partial s}$ for all a and all $s \in S$, it follows that $1-\lambda > 0$ and the higher measured wage returns to educational attainment for black workers over the range S are driven by statistical discrimination.

Under the maintained assumptions above, this proposition provides a novel way to investigate the presence of statistical discrimination operating in the labor market.

Interpreting Assumptions about Non-Home Inputs

The first assumption in the proposition above is clearly motivated by the variation of racial gaps in home inputs across education levels as documented above. A stronger sorting pattern of home inputs a into educational attainment levels s for black workers translates directly to $\bar{a}'_B(s) > \bar{a}'_W(s)$. The second assumption, however, requires further scrutiny.

This second assumption has a clear interpretation. On average, η must increase with education *at least as much* for white workers as it does for black workers. Put differently, sorting on the unobservable component η into increasingly higher levels of educational attainment cannot be stronger for black workers than for white workers. This condition is considerably weaker than the condition necessary for proper interpretation of differences in race-specific intercept terms.

With the above interpretation in mind, consider the following illustrative example. Two children, one black and one white, receive identical home inputs, and they have both attained the same level of education. If the black child faces additional barriers to skill development during this educational transition, then the marginal effect of educational attainment on productivity will be lower for this child. This is not only in line with the assumption, but there is considerable evidence that black children face additional barriers to skill development than comparable white children during the school-going years. For example, Fryer (2010) documents a constantly increasing racial test score gap that emerges at very early ages and continues to widen through twelfth grade. Fryer and Levitt (2004) document racial differences in school quality measures and their ability to explain a widening racial test score gap during the first two years of school.

Alternatively, consider a scenario in which black children face additional barriers to continued educational attainment than white children. For example, black children who go on to graduate from high school could, in principle, be unobservably higher skill than observably similar white children since they must overcome some additional barriers to high school completion. Racial differences in this type of unobserved selection into higher levels of education is therefore an important consideration.

The contribution of non-home inputs prior to the school-going years (such as neighborhood quality) must also be considered. In particular, if black high school dropouts lived in low quality neighborhoods relative to white high school dropouts with comparable early-age home inputs, it is possible that the residual racial wage gap documented for high school dropouts may be a reflection of crucial missing neighborhood inputs. Brooks-Gunn et al (1996) consider a sample of low birthweight children and find that that about half of the measured black-white gap in IQ scores at age 5 can be explained by family and neighborhood characteristics. However, after controlling for family and neighborhood characteristics as well as maternal skill measures, measures of maternal parenting behavior account for about half of the remaining gap. These results suggest a complicated interrelationship between early-age home and neighborhood influences not captured here.

The additive separability of each component of productivity in the model also deserves some discussion. For the most part, this specification is warranted given the relationships between hourly wages, HOME scores, and education in the data. However, the separability of η is less obvious, since by construction η is not observed by the econometrician. If η inputs interact with other observed inputs to productivity, then racial differences in the returns to education may reflect these interactions rather than statistical discrimination.

The remainder of this chapter uses the CNLSY79 data to provide evidence that supports the main prediction for education levels up to 13 years. The role of skill inputs from outside of the home can be investigated directly with the incorporation of several skill measures near the time of labor market entry as well as measures of neighborhood quality available in the data. The residual wage patterns described above are robust to controlling for these measures. The results suggest that, if anything, accounting for variation in these additional measures actually increases the racial difference in measured returns to education up to 13 years.

1.5 Data

The publicly available CNLSY79 data are used to construct the sample. There are four primary variables of interest constructed for each child: a race dummy, a measure of home inputs, and educational attainment measure, and early-career hourly wage measures. "Black" is an indicator assigned a value of 1 if the individual's mother identifies as black in the NLSY79 data and 0 if the individual's mother identifies as white and non-Hispanic, and the sample is restricted to these two groups. The HOME observations between ages 0 and 9 for each child are used to measure home inputs. The variable "HOME Score" is an average all the valid HOME observations, standardized and adjusted for the ages and number of observations. To measure educational attainment, highest grade completed along with receipt of GED and high school diploma are used. We restrict our attention to variation across several groups: high school dropouts who never received a GED, high school dropouts who received a GED, high school graduates, and high school graduates who attended college.[13] For the early-career hourly wage measure, weekly wage observations are divided by hours worked per week. The sample is restricted to male children only and to wages observed between ages 22 and 27.

This sample incorporates as many children as possible with sufficient data to construct measures using the variables above. In particular, the panel is restricted to children born in the years 1983-1992 since these cohorts are born early enough that the data contain some relevant early-career wage observations (which is defined as hourly wages observed at ages 22 or older) and late enough that the data have a sufficient number of HOME observations. It is worth noting that the CNLSY79 observations occur every two years and begin in the year 1986, so the 1983 and 1984 cohorts are missing exactly one HOME observation as a mechanical feature of how the data was collected. To deal with this problem along with other missing HOME observations, HOME scores are adjusted in the analysis below for the ages of observations and number of observations between ages 0 and 9.

In the CNLSY79 data, each child is linked to information about his mother through the original NLSY79 dataset. This provides access to detailed family background measures that directly affect home inputs. The home inputs are measured in the late 1980s

[13]Figures 4 and 6 in Appendix B explore how the results change when highest grade completed is used to categorize educational attainment.

Table 4: Basic Description of Panel, by Child Birth Cohort

Child Birth Cohort	Ages at HOME Observations	Ages at Wage Observations	Number of Children	Number of Wage Observations
1983	3-9	23-27	77	125
1984	2-8	22-26	96	160
1985	1-9	23-27	105	183
1986	0-8	22-26	82	147
1987	1-9	23-27	89	154
1988	0-8	22-26	105	170
1989	1-9	23-25	87	118
1990	0-8	22-24	83	114
1991	1-9	23	58	58
1992	0-8	22	45	45

and the 1990s, when shocks to black communities sustained racial differences in household resources over that period and caused a stagnation in black-white skill convergence as documented in the literature.[14] Measures of household income, age-adjusted maternal AFQT score, and marital status demonstrate the importance of racial differences in the barriers to providing adequate home environments crucial to skill development.

Finally, this chapter considers the effects of other productivity measures and inputs from outside the home. To measure skills near the time of labor market entry, the results below control for age-adjusted math and reading PIAT scores at the oldest age for each child. Variation in neighborhood quality is also captured through the use of mothers' self-reported responses about aspects of the neighborhood in the NLSY79 data.

[14] See, for example, Neal (2006) and Fryer (2010).

1.6 Results

One direct method for investigating race-specific returns to education is the estimation of a Mincerian wage equation with a few additional features. In particular, all specifications estimated in Table 5 include a race-specific intercept term. Specifications (4) and (5) also include interaction terms between race and educational attainment, while specification (4) includes an additional interaction term between race and the HOME score. The signs and magnitudes of the estimated coefficients on the race interaction terms provide simple estimates of racial differences in the wage returns to home inputs and educational attainment. Since the model only predicts race-specific returns to education up to 13 years, the sample is limited to workers with no more than 13 years of education. All specifications also control for labor market experience and year of observation.[15]

Column (1) reveals that black wages are on average about 15% lower than white wages in the sample of men. Controlling for the HOME score reduces this gap to about 10%, and controlling further for a race-independent educational attainment effect does not reduce the residual gap any further. It is clear, however, that these specifications do not take into account any race-specific returns to education that may result from statistical discrimination as discussed above. For this reason, columns (2) and (3) may not accurately capture the relationship between wages, race, and skill investments.

Imposing the restriction that racial differences in the returns to education are constant across years of education, the interaction term in column (4) suggests that the returns to each additional year of education are substantially higher for black workers

[15] As wages are only observed for workers between ages 22 and 27, years of labor market experience are relatively low across the sample. For this reason, the regression fit does not improve with the addition of higher-order experience terms or interacted experience terms as tends to be the case with longer age-earnings profiles.

than for white workers. The point estimate of this interaction term suggests that returns to each additional year of education are about 4% higher for black workers than for white workers. The wage returns to HOME score, on the other hand, do not seem to differ by race. As a result, this specification predicts that residual racial wage gap is shrinking with each additional year of education. Centering the residual black-white wage gap at 13 years of education in columns (4) and (5) reveals that the residual black-white wage gap is statistically insignificant at 13 years of education.

Table 5: Log Wage Regressions with HOME scores, Highest Grade, and Race

	(1)	(2)	(3)	(4)	(5)
Black	-0.150	-0.0960	-0.0993	-0.0517	-0.0512
	(0.0260)	(0.0260)	(0.0258)	(0.0383)	(0.0376)
HOME Score		0.0746	0.0678	0.0689	0.0670
		(0.0155)	(0.0151)	(0.0243)	(0.0151)
High Grade			0.0580	0.0400	0.0403
			(0.0145)	(0.0189)	(0.0190)
HOME*Black				-0.00356	
				(0.0302)	
Grade*Black				0.0406	0.0400
				(0.0231)	(0.0230)
Observations	721	721	721	721	721
R-squared	0.060	0.103	0.129	0.133	0.133

Standard errors in parentheses. Dependent variable is log of hourly wages between ages 22 and 27. All specifications also control for an experience term and year at wage observation. Standard errors are clustered at the individual level. Only workers with less than 14 years of education are included in the sample. In columns (4) and (5), the estimate for Black is re-centered at 13 years of education.

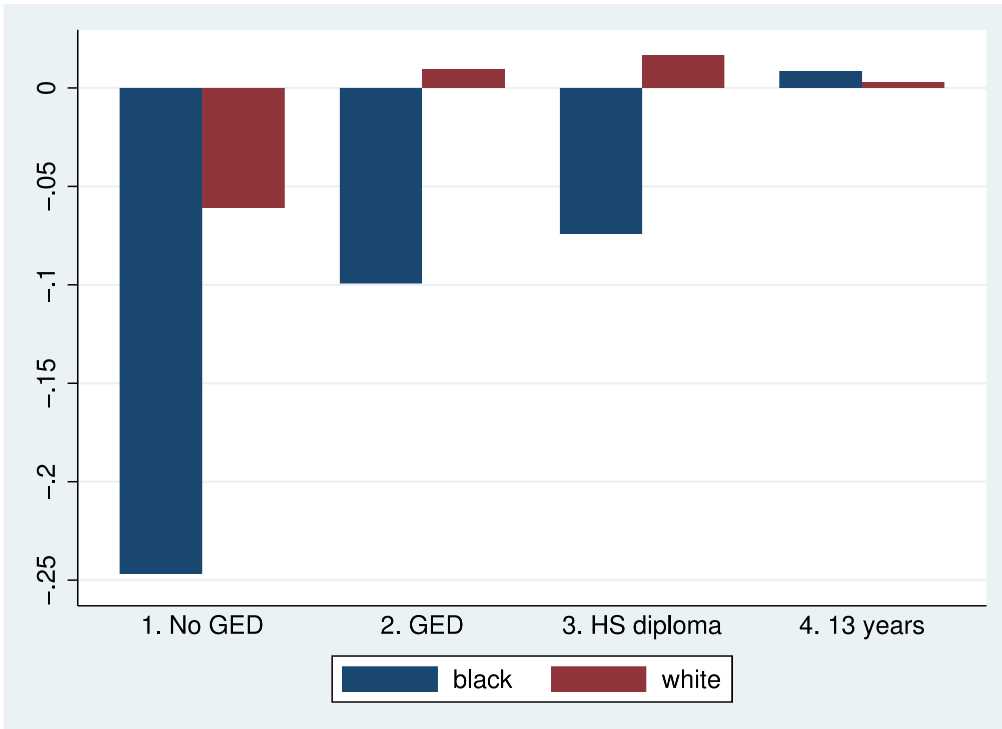
While this specification provides a basic picture of race-specific returns to education and produces one easily interpretable estimate thereof, this estimates may be misleading for a few reasons. First, the specification above makes the restrictive assumption that the racial difference in the returns to an additional year of education is the same at every level of educational attainment. Further, it seems plausible to instead consider the wage returns to increasing educational credentials (rather than simply by years of education) in this context, since this is likely what employers can observe and what they use to form expectations about worker productivity. For these reasons, a more flexible method of identifying race-specific returns to educational attainment is considered.

Figure 2a plots average log wage residuals by race and educational attainment after controlling for the HOME score, experience, and year of wage observation. For high school dropouts with no GED, the conditional racial wage gap is about 20%. For high school dropouts who received a GED, the gap shrinks to about 10%. For high school graduates, the gap is only slightly smaller. For high school diploma holders with 13 years of education, the conditional racial wage gap is essentially eliminated. This is consistent with the model's principal prediction, and it demonstrates that the racial differences in the returns to additional educational credentials are indeed large for the GED, high school diploma, and some experience in college.

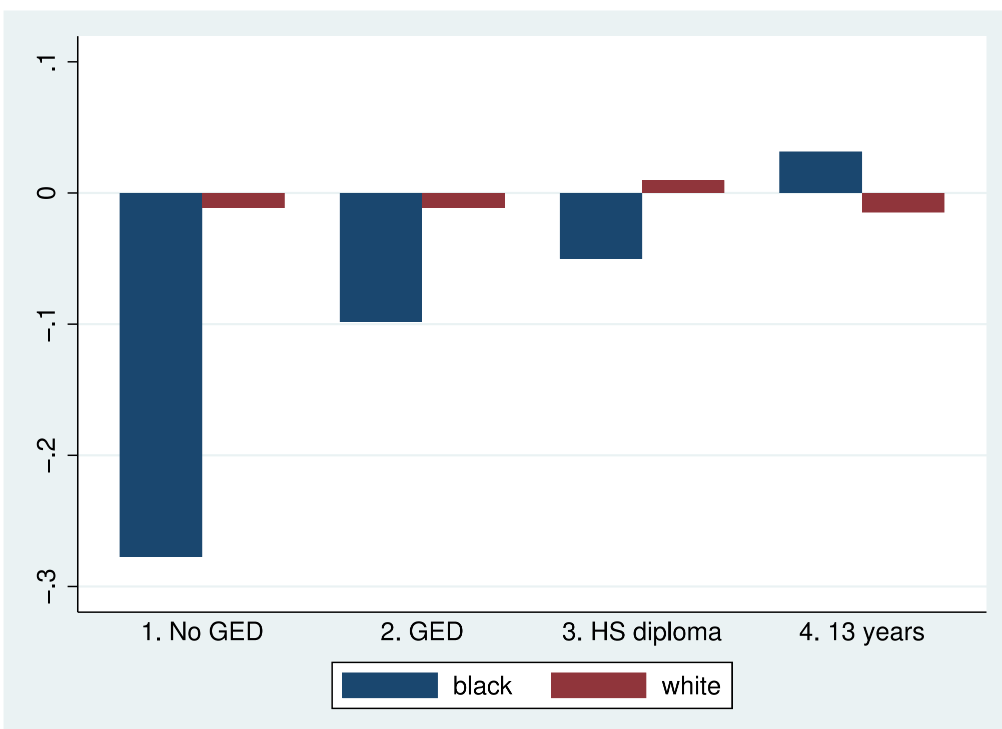
It is important to also explore the robustness of these results to the main assumption about productivity inputs that come from outside of the home. If PIAT scores near the time of labor market entry capture racial differences in skills accrued outside of the home, then consideration of how these measures can help to explain wage outcomes can provide some information about the robustness of the basic result to the main assumption. In the results presented here, the PIAT scores used are the oldest nonmissing observations for each child. These are age-adjusted, so that they provide measures of math and reading skills near the time of labor market entry that are comparable across the sample.

Figure 2: Average Adjusted Loge Wage by Race and Education

(a) Adjusted for HOME



(b) Adjusted for HOME and PIAT



(c) Adjusted for HOME, PIAT, and Neighborhood

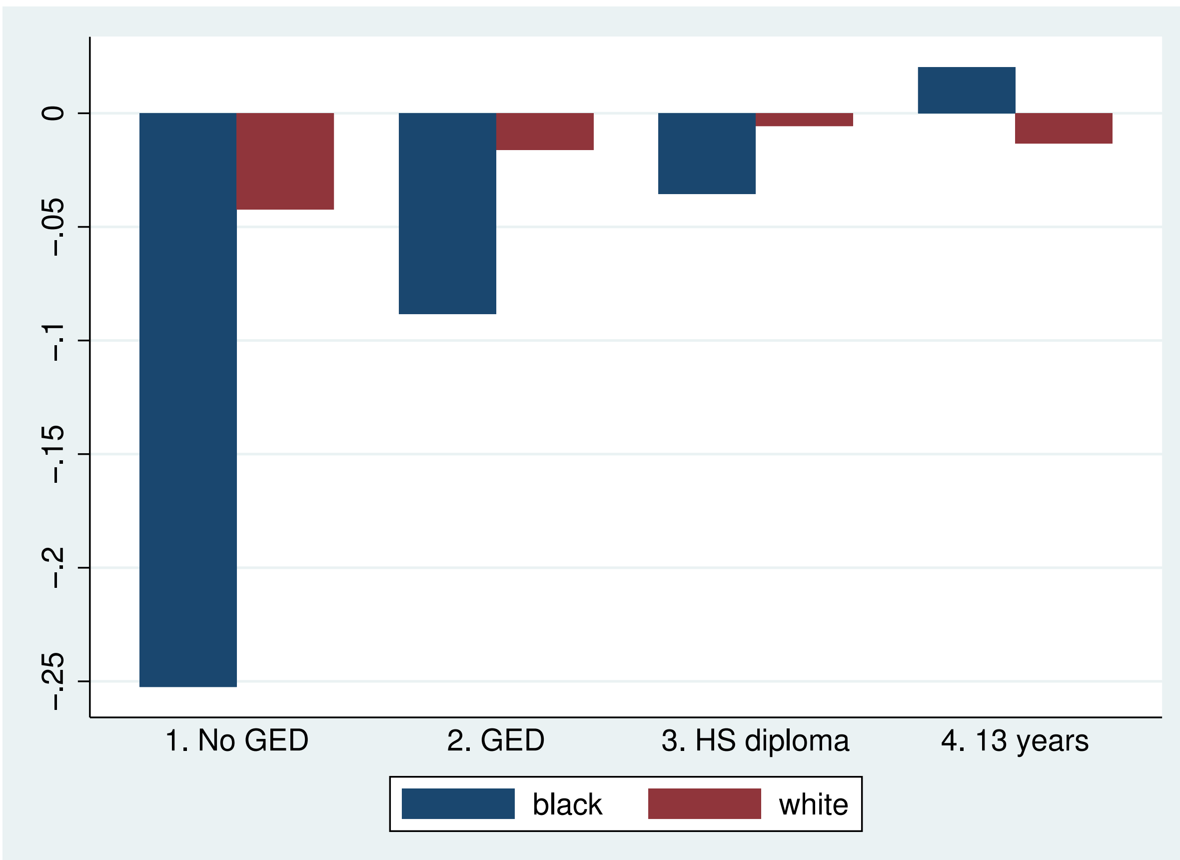


Table 6: Log Wage Regressions: Alternate Specifications with PIAT Scores

	(1)	(2)	(3)	(4)	(5)
Black	-0.150	-0.112	-0.0960	-0.0888	-0.0400
	(0.0260)	(0.0345)	(0.0260)	(0.0342)	(0.0477)
Math PIAT		0.0258		0.0234	0.0174
		(0.0217)		(0.0214)	(0.0220)
Reading PIAT		0.0160		0.00144	-0.00288
		(0.0141)		(0.0143)	(0.0144)
HOME Score			0.0746	0.0739	0.0679
			(0.0155)	(0.0182)	(0.0178)
High Grade					0.0361
					(0.0234)
Grade*Black					0.0471
					(0.0265)
Observations	721	623	721	623	623
R-squared	0.060	0.077	0.103	0.113	0.144

Standard errors in parentheses. Dependent variable is log of hourly wages observed between ages 22 and 27. All specifications also control for experience and year at wage observation. Standard errors are clustered at the individual level. Only those workers with less that 14 years of education are included in the sample. PIAT scores are the oldest nonmissing observation for each child and age-adjusted.

Table 6 explores in more detail the roles of HOME scores and PIAT scores in wage outcomes. The race-specific intercept coefficients in columns (2) through (4) reveal that HOME scores play a distinct role in explaining the black-white wage gap outside of their direct effect on PIAT scores. Further, point estimate of the coefficient on the race-education interaction term in column (5) is actually higher in this specification than in column (5) of Table 5. The inclusion of PIAT scores therefore does not substantially affect the interpretation of racial differences in measured returns to education.

Figure 2b plots average wage residuals across race and education levels after controlling for HOME score, experience, year, and a cubic polynomial term for each of the math and reading PIAT scores. Consistent with the results in Table 6, the inclusion of PIAT math and reading scores does not significantly affect the strong pattern of diminishing residual racial wage gaps with increased educational credentials. Table 6 and Figure 2b both suggest that the racial difference in returns to education persist even after the addition of PIAT scores that incorporate variation in skills acquired from non-home inputs.

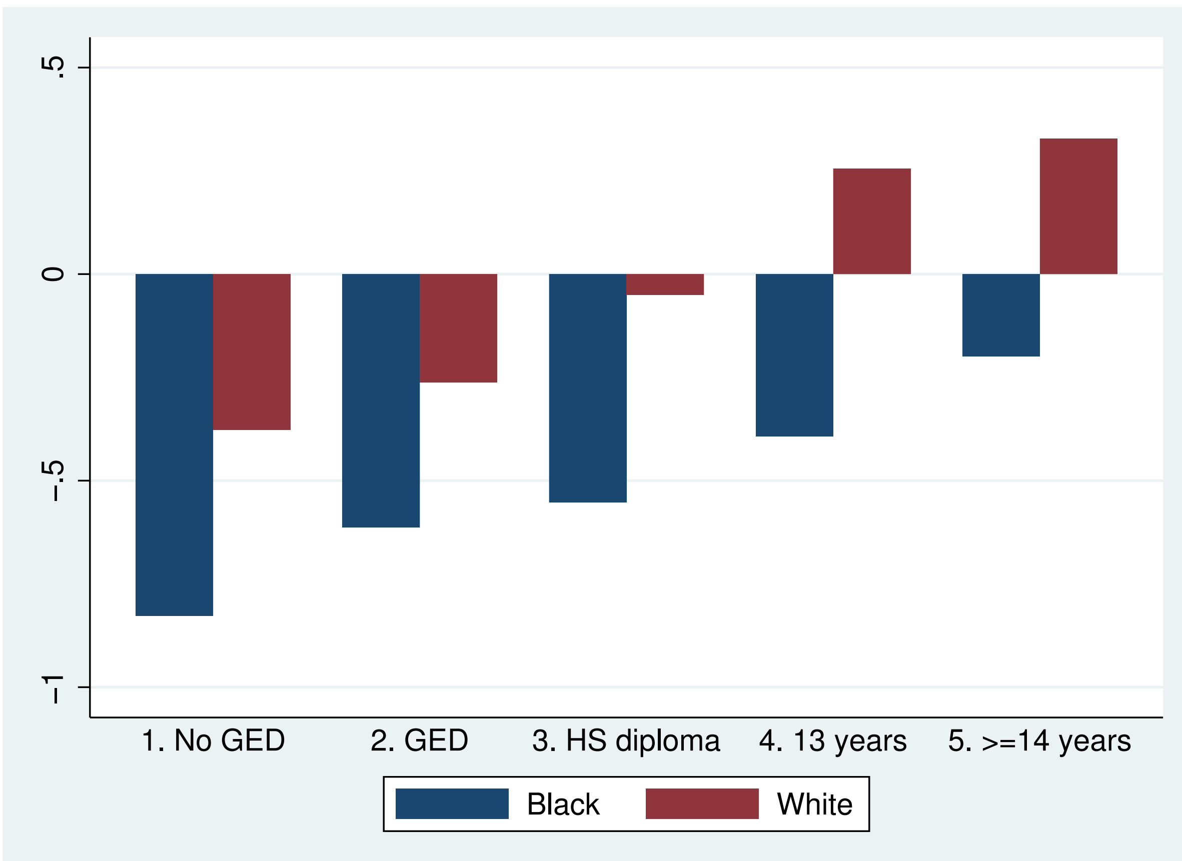
It is also useful to consider variation in measured neighborhood quality across race and education levels. These neighborhood variables may reflect crucial differences in skill development outside of the home and outside of school that are not properly captured by either the HOME score or the PIAT scores. Figure 2c also controls for several neighborhood quality measures in the NLSY79 (in addition to experience, year, HOME score, and the third-degree PIAT polynomials). The patterns are again consistent with a rapidly shrinking residual racial wage gap for educational attainment up to 13 years of education. Figure 6 in Appendix B shows a similar pattern across highest grade completed between 9 and 13 years of education.

Table 14 in Appendix B checks the robustness of Table 5 to the addition of several neighborhood quality measures as well as PIAT math and reading scores. Consistent with the results presented in this section, the estimates of the race-education interaction term are positive and large in magnitude in all specifications, and actually larger in all specifications relative to the baseline. All of these findings together support the claim that differences in important skill inputs from outside of the home cannot explain the large residual racial wage gaps for low education black workers.

Another way to check the robustness of this chapter's main prediction is to reproduce the basic results for the female respondents of the CNLSY79. If the sorting patterns of home inputs by educational attainment levels differs for black women and white women, then the model predicts that the returns to educational attainment should also differ across these groups. If these sorting patterns do not differ significantly for black and white women, then the returns to educational attainment should not differ substantially for black and white women.

Figure 3 is constructed in the same way as Figure 1 but with the female sample in the CNLSY79. In both male and female samples, higher levels of eventual educational attainment are consistently associated with higher HOME scores for both black and white women. In addition, the average HOME score is lower for black women than for white women at each level of eventual educational attainment, also consistent with patterns in the male sample. It is clear, however, that the racial gap in the HOME scores for female high school dropouts is not particularly large as it is for male high school dropouts, and the racial gap in HOME scores for women is quite constant across educational attainment levels. The sorting patterns of home inputs by educational attainment levels is quite similar for black and white women as a result.

Figure 3: Average HOME Scores by Race and Educaton: Female Sample



These patterns can be understood in the context of Autor et al (2015) and Aucejo and James (2019) which document that boys' educational attainment are more strongly predicted by family and socioeconomic disadvantage compared to girls. Further, Autor et al (2015) note that "a surprising implication of these findings is that, relative to white siblings, black boys fare worse than their sisters in significant part because black children — both boys and girls — are raised in more disadvantaged family environments." Differences in the determinants of educational attainment as well as the sensitivity to changes in these determinants for men and women, especially at the high school completion and college entry margins, can help to explain why the education sorting patterns in home inputs have a different structure for men and women.

The race-specific sorting patterns are not substantially different for black and white women, so the model predicts that the returns to education should also not differ for black and white women. Table 7 tests this prediction for women as does Table 5 for men. The results of Table 7 provides no substantial evidence that the wage returns to educational attainment differ between black and white women, since the interaction term between race and educational attainment is both close to zero and statistically insignificant.

These patterns together provide a plausible explanation for variation in the residual racial wage gap across both race and educational attainment levels. The residual racial wage gap is largest for black male high school dropouts since employer expectations of worker productivity in this group is particularly low. The residual gap shrinks with educational attainment for men since education provides a more valuable signal for black men than for white men. In contrast, employers form expectations about black women that are only modestly lower than those of their white counterparts. While these

Table 7: Log Wage Regressions: Female Sample

	(1)	(2)	(3)	(4)	(5)
Black	-0.107	-0.0725	-0.0815	-0.114	-0.104
	(0.0239)	(0.0238)	(0.0238)	(0.0322)	(0.0333)
HOME Score		0.0609	0.0564	0.110	0.0563
		(0.0156)	(0.0157)	(0.0295)	(0.0156)
High Grade			0.0477	0.0527	0.0578
			(0.0138)	(0.0189)	(0.0199)
HOME*Black				-0.0842	
				(0.0348)	
Grade*Black				-0.0149	-0.0223
				(0.0220)	(0.0221)
Observations	801	801	801	801	801
R-squared	0.046	0.070	0.089	0.101	0.091

Standard errors in parentheses

Dependent variable is log of hourly wages observed between ages 22 and 27. All specifications also control for a quadratic experience term and year at wage observation. Standard errors are clustered at the individual level. Only those workers with less that 14 years of education are included in the sample. In columns (4) and (5), the coefficient estimate for Black is re-centered at 13 years of education.

expectations rise with educational attainment, differences in the expectations for black and white women do not vary substantially across educational attainment levels. The signaling value of education is roughly same for black and white women, so the measured returns to educational attainment are not substantially for black and white women.

1.7 Discussion of Alternative Mechanisms

While the results above are consistent with the presence of statistical discrimination in the labor market as detailed in the model, it is important to address some of the model's more fundamental assumptions. First, the model maintains the assumption that the signaling noise distribution does not vary by race or by education. In the Part 1 of Appendix A, some key predictions of a more flexible model that allows differences in the signaling technology by race and education are derived. First, if the signaling technology differs by race, then the wage returns to home inputs must necessarily differ by race. This is an intuitive result, since a weaker (stronger) signal of productivity induces the firm to put less (more) weight on the signal in wage determination. However, the results from Tables 5 and 6 do not support the presence of racial differences in the wage returns to the HOME score.

The second prediction is that higher returns to education for black workers provides evidence for statistical discrimination in the labor market as long as the signaling technology becomes more clear with higher levels of education. As workers attain more education, the firm relies less on the prior belief and more on the signal. If the firm's prior is lower for black workers than for white workers at each education level, then increasing education has more value for black workers than on white workers (even when the signaling technologies do not differ by race). This is consistent with observed patterns in the HOME score.

The possibility of racial bias in the HOME score poses another problem with interpreting these results. This type of racial bias has been discussed in several other studies that link racial wage gaps to racial gaps in measured skills. In particular, Neal and Johnson (1996) cite an external study of racial bias in the AFQT as a predictor of measured military job performance and use this study as evidence that the AFQT is unbiased as a measure of labor market productivity. However, Altonji and Blank (1999) question the generalizability of this study to civilian job performance. They also point out that varying racial differences in the wage returns to of the AFQT differs by components, with the verbal score mattering more for black workers. This difference is consistent with the presence of racial bias in the AFQT, as documented by Rodgers and Spriggs (1996).

Part 2 of Appendix A explicitly incorporates systematic racial bias in the HOME score measurement into the model. The extended result makes it clear that in order for racial bias alone to drive the racial difference in returns to education, it must be the case that racial bias has the smallest effect for the lowest education workers and rises with educational attainment. In contrast, racial bias can only explain the narrowing racial gap in home inputs if this bias has the largest effect for the lowest education workers. Racial bias could, in principle, explain one of these patterns. However, these results show that racial bias alone cannot generate both patterns.

Next, the model above does not incorporate censoring due to labor market participation. Restricting the analysis to only observed wages conditional on labor market participation may not accurately depict racial differences in wage offer distributions. Neal and

Johnson (1996) note that there are substantial problems with pursuing a structural approach to modeling labor market participation because of the lack of a reliable exclusion restriction for male labor market participation. Johnson et al (2000) demonstrate that imputing a log wage of zero for every nonparticipant and running a median wage regression provides consistent estimates for the wage offer distribution under the assumption that wage offers of nonparticipants are below the median wages of observably similar participants. While this imputation method requires some strong assumptions about the wage offers of nonparticipants, it is useful as a tool for understanding the effect of censoring on the regression results in the case that nonparticipants receive offers less than the median wage in the population conditional on their observable characteristics.

In Table 8, such a median wage regression is estimated. The results are qualitatively similar to those in Table 5, but there are some notable differences. Since the black nonparticipation rate is higher than the white nonparticipation rate, the estimated residual racial wage gap from the median regression is larger the estimate from the ordinary least-squares regression. Further, the racial difference in nonparticipation rates is largest at the lowest levels of education and narrows with educational attainment.

Statistical discrimination is likely to affect outcomes at both the intensive wage margin and the extensive participation margin. In this context, the median wage regression captures both effects while the ordinary least-squares wage regression will only pick up effects on the intensive margin. With this in mind, the ordinary least-squares regression estimates will tend to understate the racial difference in the labor market returns to education. The estimates from the median regression help to confirm the importance of this intuition. In column (4) and (5) of Table 8, the coefficient estimates on the education interaction term is both larger and more significant than the coefficient estimates on the education interaction term from Table 5.

Table 8: Median Wage Regressions with HOME scores, Highest Grade, and Race

	(1)	(2)	(3)	(4)	(5)
Black	-0.171	-0.0935	-0.119	-0.0621	-0.0569
	(0.0272)	(0.0267)	(0.0297)	(0.0421)	(0.0409)
HOME Score		0.0891	0.0806	0.0814	0.0786
		(0.0155)	(0.0172)	(0.0243)	(0.0162)
High Grade			0.0659	0.0506	0.0506
			(0.0158)	(0.0175)	(0.0172)
HOME*Black				-0.00872	
				(0.0327)	
Grade*Black				0.0539	0.0514
				(0.0241)	(0.0234)
Observations	935	935	935	935	935

Dependent variable is log of hourly wages observed between ages 22 and 27. All specifications also control for an experience term and year at wage observation. Only those workers with less that 14 years of education are included in the sample. Nonparticipants with no wage observations are assigned an hourly wage of one dollar. In columns (4) and (5), the coefficient estimate for Black is re-centered at 13 years of education.

The model also does not incorporate search frictions. Fryer et al (2013) take advantage of an extensive dataset of unemployed workers in New Jersey to measure racial differences in job search strategies and outcomes. They find that, conditional on previous wage, there are few important differences in search strategies outside of differences in reservation wages. Offered wages to black workers are also lower that offered wages to white workers, conditional on previous wage and other factors. This suggests that search frictions may tend to exacerbate racial disparities in the labor market that arise from statistical discrimination on the basis of race.

1.8 A Few Policy Considerations

Much of the literature on mitigating racial disparities in labor market and other adult outcomes suggest that eliminating barriers to early-age skill development, especially among the children who grow up in the most under-resourced households, constitutes an important policy goal. In light of the results above, this goal should certainly include improving early-age home environments. The results also point to a new policy consideration. If a high-quality intervention can be successfully implemented on a large enough scale, improvements would not only have the obvious direct effect on individual skill development. Improvements across the population of low education black workers can additionally improve employer beliefs about these workers' productivity levels, further mitigating racial inequality in the labor market.

Cunha et al (2015) demonstrate that in a sample of disadvantaged mothers, subjective beliefs about the importance of early-age home environments do not match objective measured skill benefits of these environments. A large-scale information campaign that teaches mothers about the importance of the home environment as well as instruction and support in providing these environments might therefore be particularly effective. While this area of research shows much promise, the desired program might be especially difficult to implement effectively for the most important populations, as these information campaigns can differentially affect members of different socioeconomic groups.[16]

If early-age home and preschool environments are substitutes, the results from the Perry Preschool Program and other similar projects[17] suggest one possible alternative.

[16]See, for example, Aizer and Stroud (2010) for evidence of this differential impact in a related example surrounding the release of the 1964 Surgeon General's Report on Smoking and Health.

[17]Several other studies of intensive preschool interventions such as the Carolina Abecedarian Project have found positive results on eventual adult outcomes similar to those of the Perry Preschool Project.

Between 1962 and 1967, the Perry Preschool Program provided a high-quality preschool education to children ages 3 and 4. The program provided treatment for a sample of African-American children living in poverty in a randomized control trial setting. Comparing treatment and control groups through age 40, results from the study demonstrate that the effect of treatment on both labor market outcomes and adult crime are substantially large. While the effect of treatment on IQ at age 5 is large, these IQ differences degrade rapidly over time, and are no longer statistically significant at any time beyond one year after treatment.[18] While the results presented in this chapter concern variation in home inputs rather than preschool inputs, the substitutability of early-age preschool and home environments constitutes an important area of future study.

1.9 Conclusion

While a large portion of the racial wage gap can be attributed to racial differences in skill bundles at the time of labor market entry, evidence presented in this chapter suggests that part of the racial wage gap for low education workers may reflect the operation of statistical discrimination in the labor market. Particularly large racial gaps in early-age home productivity inputs of low education workers can drive employers to resort to statistical discrimination in wage assignment. The evidence suggests that policy interventions should focus primarily on expanding the scope of early-age intervention programs, especially for children living in poverty.

[18]See Schweinhart et al (2005) for more details.

Bibliography

[1] Aiger, D. J. and Cain, G. G. (1977). Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review*, 30(2):175-187.

[2] Aizer, A. and Stroud, L. (2010). Education, Knowledge and the Evolution of Disparities in Health. Working Paper No.15840, National Bureau of Economic Research, Inc.

[3] Altonji, J. G. and Blank, R. M. (1999). Race and Gender in the Labor Market. *Handbook of Labor Economics*, 3:3143-3259.

[4] Altonji, J. G. and Pierret, C. R. (2001). Employer Learning and Statistical Discrimination. *Quarterly Journal of Economics*, 116(1):313-350.

[5] Arcidiacono, P., Bayer, P., and Hizmo, A. (2010). Beyond Signaling and Human Capital: Education and the Revelation of Ability. *American Economic Journal: Applied Economics*, 2:76-104.

[6] Arrow, K. J. (1973). *The Theory of Discrimination.* In Discrimination in Labor Markets. Princeton University Press.

[7] Aucejo, E. M. and James, J. (2019). Catching up to girls: Understanding the gender imbalance in educational attainment within race. *Journal of Applied Econometrics.*

[8] Autor, D., Figlio, W., Karbownik, K., Roth, J., and Wasserman, M. (forthcoming). Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes. *American Economic Journal: Applied Economics.*

[9] Bowlus, A. J. and Eckstein, Z. (2002). Discrimination and Skill Differences in an Equilibrium Search Model. *International Economic Review*, 43(4):1309-1345.

[10] Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4):991-1013.

[11] Brooks-Gunn, J., Klebanov, P. K., and Duncan, G. J. (1996). Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment, and Maternal Characteristics. *Child Development*, 67(2):396-408.

[12] Brooks-Gunn, J., and Markman, L. (2005). The contribution of parenting to ethnic and racial gaps in school readiness. *The Future of Children*, 15(1):139-168.

[13] Caldwell, B. and Bradley, R. (1984). *Home observation for measurement of the environment.* University of Arkansas Press.

[14] Carneiro, P., Heckman, J. J., and Masterov, D. V. (2005). Understanding the Sources of Ethnic and Racial Wage Gaps and Their Implications for Policy. *Handbook of Employment Discrimination Research*, 99(136):99-136.

[15] Coate, S. and Loury, G. C. (1993). Will Affirmative-Action Policies Eliminate Negative Stereotypes? *American Economic Review*, 83(5):1220-1240.

[16] Cunha, F., Elo, I., and Culhane, J. (2013). Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation. Working Paper No. 19144, National Bureau of Economic Research, Inc.

[17] Cunha, F., Heckman J. J., and Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3):883-931

[18] Ferguson, R. (2005). *America's black-white school achievement gap persists.* In Ethnicity, Social Mobility, and Public Policy: Comparing the USA and UK. Cambridge University Press.

[19] Fryer, R. G. (2010). Racial Inequality in the 21st Century: The Declining Significance of Discrimination. *Handbook of Labor Economics*, 4:855-972.

[20] Fryer, R. G. and Levitt S. D. (2004). Understanding the Black-White Test Score Gap in the First Two Years of School. *The Review of Economics and Statistics*, 86(2):447-464.

[21] Fryer, R. G., Pager D., and Spenkuch, J. L. (2013). Racial Disparities in Job Finding and Offered Wages. *Journal of Law and Economics*, 56:633-689.

[22] Johnson, W., Kitamura, Y., and Neal, D. (2000). Evaluating a Simple Method for Estimating Black-White Gaps in Median Wages. *American Economic Review*, 90(2): 339-343.

[23] Kan, K., and Tsai, W. (2005). Parenting practices and children's education outcomes. *Economics of Education Review*, 24(1):29-43.

[24] Lang, K. and Manove M. (2011). Education and Labor Market Discrimination. *American Economic Review*, 101(4):1467-1496.

[25] Lundberg, S. J. and Startz, R. (1983). Private Discrimination and Social Intervention in Competitive Labor Market. *American Economic Review*, 73(3):340-347.

[26] Neal, D. A. (2006). Why has Black-White Skill Convergence Stopped? *Handbook of the Economics of Education*, 1:511-576.

[27] Neal, D. A. and Johnson, W. R. (1996). The Role of Premarket Factors in Black-White Wage Differences. *Journal of Political Economy*, 104(5):869-895.

[28] Rodgers, W. and Spriggs W. (1996). What does the AFQT really measure: race, wages, schooling and the AFQT score. *Review of Black Political Economy*, 24(4):13-46.

[29] Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., and Nores, M. (2005). Lifetime effects: The High/Scope Perry Preschool study through age 40. (Monographs of the High/Scope Educational Research Foundation, 14). Ypsilanti, MI: High/Scope Press.

[30] Thompson, Owen, (2018). The Determinants of Racial Differences in Parenting Practices. *Journal of Political Economy*, 126(1):438-449.

[31] Todd, P. E., and Wolpin K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, 113:F3-F33.

[32] Todd, P. E., and Wolpin K. I. (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, 1(1):91-136.

Chapter 2

Skill Development, Early-Career Occupational Sorting, and College Entry Timing

2.1 Introduction

Approximately one-third of college enrollees in the United States wait at least one year after high school graduation before first enrolling in college according to Wine et al (2013). Despite the prevalence of delayed college entry, relatively few studies examine the consequences of delayed college entry for future labor market outcomes.

This chapter finds that, on average, delayed college entrants have lower earnings trajectories than observably similar immediate college entrants. These results are consistent with estimates from a similar comparison in Lin and Liu (2019), and Light (1995) documents a future wage penalty to more general interrupted schooling patterns. Light (1995) argues that these earnings differences may stem from the fact that skills developed through work experience and college can depend not only upon the quality of these investments but also on the order in which these investments are made. In this context, delaying college entry and gaining work experience beforehand can have different implications for skill development than entering college immediately following high school graduation.

The principal goal of this chapter is to explore how differences in early-career occupation choices affect the dynamic interaction of skill investments both in college and in the early-career. In particular, certain types of occupational investments may complement college best when performed before college entry, while other occupational investments may be more productive after college completion. If this is the case, then the optimal occupation choice of a delayed college entrant between high school and college would likely differ from the optimal occupation choice of an immediate college entrant directly following college completion. This can also create differences in the penalty associated with delaying college entry for individuals with different levels of skills at the time of high school graduation. These basic insights can help shed light on the observed relationships between early-career occupation choices, college entry timing choices, and future earnings.

A simple learning-by-doing framework is used to interpret variation in future income trajectories. Data from the National Longitudinal Survey of Youth 1979 Cohort (NLSY79) are used to create a sample of college graduates, including individual skill measures at the time of high school graduation, timing of college entry, and early-career occupation choices. Occupations are also linked to both abstract and routine occupational task content data. Together, these data are used to construct measures of each worker's speed of abstract and routine task completion at the time of high school graduation and number of abstract and routine tasks completed in the early career. Each plays a critical role in the development of abstract and routine skills through college and the early career. The model allows for differences in skill growth across college entry timing decisions to capture an important feature of the dynamic skill development framework.

These data are used to explore how early-career occupation choices affect future earnings, and how these relationships differ across college entry timing decisions. Several important relationships are documented. In general, more abstract tasks and more routine tasks completed in the early career are associated with higher income trajectories, and abstract task completion generally has a larger impact on future income than routine task completion. However, the benefits of early-career task completion are not the same for immediate and delayed college entrants. Immediate entrants enjoy larger benefits from abstract task completion, while delayed college entrants enjoy larger benefits from routine task completion. These patterns suggest that abstract skill investments in the early career complement college most when performed before college, whereas routine skill investments in the early career are most effective when performed after college.

In order to maximize the skill benefits of early-career labor force experience, each worker should choose a more routine-intensive occupation if they delay college entry but a more abstract-intensive occupation after college completion if they immediately enter college. Do workers account for these incentives when choosing early-career occupations? Comparing the early-career occupation choices of delayed and immediate college reveals that immediate college entrants choose relatively abstract-intensive early-career occupations when compared to observably similar delayed college entrants. This provides evidence that the early-career occupations are chosen, at least in part, to best complement skill development in college.

The last part of this chapter explores the contributions of several important determinants of college entry timing decisions. High school graduates with higher abstract test scores are significantly less likely to delay college entry. This remains true after accounting for differences in demographic, family background, and noncognitive assessment measures. In contrast, differences in routine test scores do not significantly impact the timing of college entry. The dynamic skill development framework provides a plausible explanation for these patterns, as individuals with high abstract test scores face the largest penalty to working after high school graduation rather than continuing skill development in college. This can inform policymakers who seek to minimize the harmful effects of delayed college entry on future labor market outcomes. In particular, policies that encourage immediate college entry should target individuals with both high abstract skills and other background characteristics that create additional barriers to immediate college entry.

The chapter proceeds as follows. Section 2 briefly describes some related literature. Section 3 develops a model of college entry timing and early-career occupation choices. Section 4 describes the data used, and Sections 5 through 7 present the main results. Section 8 concludes.

2.2 Related Literature

In the literature, several possible reasons for delayed college entry are identified. Bozick and DeLuca (2005) demonstrate that for some high school graduates, military service, sickness, marriage, pregnancy, or a death in the family can affect college entry timing decisions. Horn et al (2005) documents that in recent years, taking a "gap year" has become more commonplace in the United States, especially for students already admitted to elite undergraduate institutions. However, as noted by Goldrick-Rab and Han (2011), most delayed college entrants do not fit this explanation and other explanations are likely to play a more crucial role in the decision to delay for these individuals.

Some studies have examined tuition and borrowing constraints as possible reasons for delayed entry, as high school graduates who delay college entry can work and save up for future tuition payments. For example, Kane (1996) makes the case that binding borrowing constraints are an important deterrent to college entry and continual college enrollment. In contrast, Carneiro and Heckman (2002) and Cameron and Taber (2004) find little evidence that borrowing constraints affect college enrollment decisions at the margin. Johnson (2013) considers delayed college entry decisions in particular and estimates a relatively small impact of borrowing constraints on college entry timing decisions.

Lin and Liu (2019) also compare the outcomes of delayed and immediate college enrollees and find a significant penalty to delayed college enrollment. While selection on unobserved characteristics may be driving some of these differences in future earnings, Lin and Liu (2019) provide evidence that the income penalty to delayed college enrollment is relatively robust to selection of this type when assumption that selection on unobservables is assumed to be proportional to selection on observables. Griliches (1980) and Light (1995) also study the effect of the order in which schooling and work are accumulated on income. Griliches (1980) finds no evidence of a significant effect while Light (1995) estimates a penalty to the work experience accumulated before finishing college using more recent data. All of these studies compare individuals with the same schooling and years of work experience but ignore the effects of differences in the early-career occupations that are central to this chapter.

While this chapter focuses on the ramifications of delayed college entry, there is a growing literature that seeks to understand “college stopout,” or taking time off from college. Arcidiacono et al (2016) and Pugatch (2018) explore how individuals learn about college and labor market ability through exposure to both college and the labor force, and Gaulke (2014) considers the role of credit constraints in the stopout decision. Goldrick-Rab (2006) demonstrates that students with interrupted schooling patterns disproportionately come from lower socioeconomic backgrounds.

This chapter is also related to a literature that uses occupational task-specific human capital to explain observed differences in wage growth across occupations. Yamaguchi (2012) estimates a structural model of multidimensional human capital in which skill evolution depends upon the bundle of occupational tasks chosen. Sanders (2016) considers the importance of skill accumulation relative to skill uncertainty in occupational transitions over the life cycle. Lise and Postel-Vinay (2015) embed task-specific human capital growth into a job search model. None of these papers explicitly examine the interaction of skill accumulation through occupational task completion and skill accumulation through college attendance.

2.3 A Model of Skill Development, College Entry Timing, and Early-Career Occupation Choices with Two Types of Skills

The following model illustrates how dynamic skill development can affect the incentives that drive college entry timing and early-career occupation choices. It will also help to clarify how these incentives contribute to the relationships between college entry timing, early-career occupation choices, and future earnings trajectories observed in the data.

At the time of high school graduation, each individual is endowed with a speed of abstract task completion, $h_i^A \geq 0$, and a speed of routine task completion, $h_i^R \geq 0$. Each individual chooses to either enter college immediately or delay college entry, a choice denoted by $e \in \{I, D\}$. In either case, each individual also chooses an early-career occupation. Occupations are characterized by the vector (X^A, X^R), where $X^A \geq 0$ is the amount of time spent on abstract tasks and $X^R \geq 0$ is the amount of time spent on routine tasks in the early-career.

After skill development through college and the early-career, the improved abstract and routine skill speeds are represented by $F^A(h^A, X^A, e)$ and $F^R(h^R, X^R, e)$, respectively. Assume that these skill development processes take the following form.

$$F^s(h^s, X^s, e) = h^s + \kappa_{se} h^s X^s$$

Skill development, or speed improvement, depends crucially on the number of abstract and routine tasks completed in the early-career. Note that the number of abstract tasks is the product $h^A X^A$ and the number of routine tasks by $h^R X^R$. Completion of abstract and routine tasks may affect skills differently depending upon whether they are performed before college entry or after college completion. This form introduces the restricting assumption that the speed of task completion increases at a constant rate κ_{se} for every task of type $s \in \{A, R\}$ performed in the early-career with college entry timing e.

Further, individuals face common prices P^A and P^R for abstract and routine skill speeds in the labor market after college and the early-career. That is, an individual

i endowed with speeds (h_i^A, h_i^R) and makes entry timing and early-career occupation choices (e, X^A, X^R) earns the following wages after college and the early-career.

$$w_i(e, X^A, X^R) = P^A F^A(h_i^A, X^A, e) + P^R F^R(h_i^R, X^R, e)$$

The principal objective in the early career is skill development, while the principal objective in the late career is income maximization. Individual i chooses (e, X^A, X^R) to maximize utility, which depends upon both future wages and costs of abstract and routine time investments in the early-career.[1]

$$u_i(e, X^A, X^R) = w_i(e, X^A, X^R) - c_A(X^A) - c_R(X^R)$$

For each possible college entry timing choice, the associated optimal early-career occupation choice is the solution to the following optimization problem.

$$(X_i^A(e), X_i^R(e)) = \underset{(X^A, X^R)}{\text{argmax}}\{u_i(e, X^A, X^R)\}$$

For each of the two skill types $s \in \{A, R\}$, a first-order condition can be derived.

$$P^s h_i^s \kappa_{se} = c_s'(X_i^s(e))$$

Assume that the cost functions satisfy $c_s'(0) = 0$, $c_s'' > 0$, and $\lim_{X \to \infty} c_s'(X) = \infty$.

[1] It should be noted that utility costs for time spent on routine and abstract tasks are additively separable. This departs from the standard model of labor-leisure choice in which the optimal hours worked derives from some tradeoff between leisure and consumption goods. However, separable tastes for different occupational task components is a common feature of other task-specific human capital models, including Yamaguchi (2012) and Lise and Postely-Vinay (2015).

Under these assumptions, there exists a unique optimal early-career occupation choice $(X_i^A(e), X_i^R(e))$ for each choice of college entry timing e.

A few results follow immediately.

1. *Fixing h_i^R, a higher (lower) h_i^A will choose a higher (lower) $X_i^A(e)$.*

2. *Fixing h_i^A, a higher (lower) h_i^R will choose a higher (lower) $X_i^R(e)$.*

3. *$X_i^A(e = I) > (<) X_i^A(e = D)$ if $\kappa_{IA} > (<) \kappa_{DA}$.*

4. *$X_i^R(e = I) > (<) X_i^R(e = D)$ if $\kappa_{IR} > (<) \kappa_{DR}$.*

Statements 1 and 2 concern differences optimal early-career occupation choice for individuals with different skill endowments. All else equal, individuals with higher endowments of abstract skills will invest more in abstract tasks in the early career. An analogous statement is true for routine skills. These statements hold for each college entry timing decision.

Statements 3 and 4 concern the different occupation choices for individuals with the same skill endowments but different choices of college entry timing. Immediate college entrants will invest more in abstract skills than their delayed entrant counterparts if and only if they enjoy relatively higher returns to abstract tasks completion in the early career. An analogous statement is true for routine skills.

Do individuals with higher abstract (routine) skill endowments choose more abstract-intensive (routine-intensive) early-career occupations? Do the future income benefits of abstract and routine task completion differ by college entry timing? If so, are differences in the early-career occupation choices between immediate college entrants and delayed college entrants consistent with these differences in returns? All of these questions are explored directly in the empirical results presented in this chapter.

Let us now explore the consequences of delaying college entry. Given the optimal early-career occupation choices, each individual chooses the college entry timing e_i that maximizes utility.

$$e_i = \underset{e \in \{I,D\}}{\text{argmax}} \{u_i(e, X_i^A(e), X_i^R(e))\}$$

Defining $W_i(e) = w_i(e, X_i^A(e), X_i^R(e))$ and $C_i(e) = c_A(X_i^A(e)) + c_R(X_i^R(e))$, the optimal college entry timing e_i has the following threshold rule.

$$e_i = \begin{cases} I & \text{if } [W_i(I) - W_i(D)] - [C_i(I) - C_i(D)] > 0 \\ D & \text{otherwise} \end{cases}$$

The two determinants of the timing decision are the magnitude of the future income penalty to delayed college entry and the difference in the utility cost of time spent on both abstract and manual tasks in the early career. Because individuals differ in their skill endowments, each of these factors may also differ. This can result in differences incentives to delay college entry for individuals with different skill endowments.

In order to examine this more closely, assume that the skill development function has the following form, the future income penalty to delayed college entry can be derived.

$$W_i(I) - W_i(D) = \sum_{s=A,R} P^s h_i^s [\kappa_{sI} X_i^s(I) - \kappa_{sD} X_i^s(D)]$$

For each skill, the future income penalty to delayed college entry depends upon differences in the number of tasks completed as well as differences in the returns to those tasks across immediate and delayed entry choices.

What is the relationship between skill endowments and the future income penalty to delay? Fixing h_i^R, consider an increase in the endowed h_i^A. There are three relevant cases.

First, consider $\kappa_{AI} = \kappa_{AD}$. The first-order condition for the early-career occupation choice implies that $X_i^A(I) = X_i^A(D)$ for any value of h_i^A. In this case, $W_i(I) - W_i(D)$ will not change as a result of the increase in h_i^A.

Next, consider $\kappa_{AI} > \kappa_{AD}$. Since c_A is convex, the first-order condition for the early-career occupation choice implies that $X_i^A(I)$ will increase *more* than $X_i^A(D)$ as a result of the increase in h_i^A. This implies that $W_i(I) - W_i(D)$ will increase as a result of the increase in h_i^A.

Finally, consider $\kappa_{AI} < \kappa_{AD}$. Now, the first-order condition for the early-career occupation choice implies that $X_i^A(I)$ will increase *less* than $X_i^A(D)$ as a result of the increase in h_i^A. This implies that $W_i(I) - W_i(D)$ will decrease as a result of the increase in h_i^A.

A symmetric argument reveals how changes in h_i^R affect the future income penalty to delayed entry $W_i(I) - W_i(D)$. The results are summarized as Statements 5 through 8.

5. *Fixing h_i^R, a higher (lower) h_i^A has a larger (smaller) future income penalty to delayed entry if $\kappa_{AI} > \kappa_{AD}$.*

6. *Fixing h_i^R, a higher (lower) h_i^A has a smaller (larger) future income penalty to delayed entry if $\kappa_{AI} < \kappa_{AD}$.*

7. *Fixing h_i^A, a higher (lower) h_i^R has a larger (smaller) future income penalty to*

delayed entry if $\kappa_{RI} > \kappa_{RD}$.

8. *Fixing* h_i^A, *a higher (lower)* h_i^R *has a smaller (larger) future income penalty to delayed entry if* $\kappa_{RI} < \kappa_{RD}$.

Note that for any change in skill endowment h_i^s, $W_i(I) - W_i(D)$ and $C_i(I) - C_i(D)$ will move in the same direction. Let us focus on the following scenario. For any change in h_i^s, the associated change in $C_i(I)-C_i(D)$ is smaller than the change in $W_i(I)-W_i(D)$.[2] Considering that the income portion represents the total benefit of future lifetime income streams, this scenario seems quite plausible. Now, consider an increase in h_i^A when $\kappa_{AI} > \kappa_{AD}$. For a higher h_i^A, the increased future wage penalty for delayed college entry outweighs the higher additional cost of investing more in X_i^A as an immediate college entrant. This also implies that, all else equal, higher h_i^s individuals will be *more* inclined to enter college immediately when $\kappa_{sI} > \kappa_{sD}$ and *less* inclined when $\kappa_{sI} < \kappa_{sD}$. This provides an important insight into how the likelihood of delaying college varies across individuals with different skill endowments that will be investigated in the empirical results to follow.

2.4 Data

The empirical analyses in this chapter use a sample constructed from two sources: the publicly available NLSY79 data and occupational task content measures constructed by Autor and Dorn (2013).[3] The sample consists of all individuals in the NLSY79 with: (1) a high school diploma by age 19; (2) reported occupation choices immediately after

[2]It is straightforward to show that a specification with quadratic cost functions is consistent with this scenario.

[3]These data are available for use from the author's website http://www.ddorn.net/data.htm.

entering the labor force; (3) a four-year college degree at the time 10 years of labor force experience have been accumulated; and (4) reported incomes at least 10 years of labor force experience have been accumulated. This sample can be used to meaningfully compare immediate enrollees and delayed enrollees. Income observations after 10 years of labor force are used to directly compare the future labor market outcomes immediate enrollees and delayed enrollees with the same level of labor force experience.

A binary variable $D_i \in \{0, 1\}$ determines whether individual i enrolled immediately in college after high school graduation or delayed college enrollment. This is determined by comparing the year of high school graduation to the first year of college enrollment observed. If data on college enrollment are missing immediately after the time of high school graduation, the individual is dropped from the panel. The final sample consists of 1,782 individuals, 1,296 of whom are immediate college enrollees $D_i = 1$ and 486 of whom are delayed college enrollees $D_i = 0$.

The assumption is made that individual either choose college enrollment or labor force participation in each year after high school graduation. For each year that an individual is not enrolled in college, the individual gains a year of labor force experience, totaling to exp_{it}. Log incomes w_{it} of individual i in year t are recorded for all observations in which $exp_{it} \geq 10$. The total number of wage observations in the final sample is 21,143.

This chapter uses the ASVAB (Armed Services Vocational Aptitude Battery) to construct measures of abstract and routine skills at the time of high school graduation. First, the fraction of correct answers on each of six different tests (Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Mathematics Knowledge, Coding, and Numerical Operations) are observed for each individual. All participants face a time limit for each test. If it is assumed that the time limits on these tests are binding constraints for each individual, then we can think of the raw scores as measures of the speed of task completion in each subject.

To consider how these tests measure abstract and routine skills, principal component analysis is used to convert the six test score percentages into two latent variables. A clear pattern emerges. The first four tests factor into one latent variable, while the latter two tests factor into the other latent variable. Given the abstract and routine task interpretation, it is assumed that the first four tests (Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Mathematics Knowledge) measure abstract task completion speed while the latter two tests (Coding and Numerical Operations) measure routine task completion. These two principal components are interpreted as the speed of abstract and routine task completion, respectively. Finally, these components undergo a linear transformation so that they have the same mean and variance but the interpretation of task completion speed is preserved. The resulting abstract and routine skill measures are referred to as H_i^A and H_i^R, respectively.

For each individual, the early-career occupation is defined as the first observed occupation after high school graduation and before college entry for delayed college entrants and the first occupation after college graduation immediate college entrants. Using a crosswalk to the 1970 Census occupation codes, the occupations in the NLSY79 are merged with corresponding occupational task content measures constructed by Autor and Dorn (2013) from the Dictionary of Occupational Titles. Routine and abstract occupational task contents correspond to routine and abstract skill speeds constructed from the ASVAB. For each individual i, abstract task content X_i^A measures the importance of math and reading skills as well as directing, controlling, and planning activities of others in early-career occupational tasks. Routine task content X_i^R measures the importance of attaining precise set limits, tolerances, and standards in the early-career occupational tasks. These task measures are taken to represent the amount of time dedicated to abstract and manual tasks in each occupation.

With measures H_i^A and H_i^R of the speed of abstract and routine task completion and measures X_i^A and X_i^R of the amount of time spent on abstract and routine task completion in the early career, simple multiplication gives measures of the number of abstract and routine tasks completed in the early career. These are $T_i^A = H_i^A X_i^A$ and $T_i^R = H_i^R X_i^R$, respectively. Given the learning-by-doing specification, T_i^A and T_i^R are crucial to skill development and therefore included as important determinants of future income trajectories in the analyses to follow.

Table 9: Summary Statistics, by Enrollment Timing

	All mean	Immediate mean	Delay mean
Male	44%	45%	40%
Black	22%	22%	22%
Abstract ASVAB	0.67	0.75	0.36
Routine ASVAB	0.50	0.53	0.37
Abstract Task Content	3.29	3.32	3.15
Routine Task Content	3.86	3.77	4.17
Family Income/1000	22.8	24.2	18.0
Parent Went to College	48%	53%	32%
Rotter Score	8.13	8.04	8.48
Rosenberg Score	23.5	23.7	23.0
Observations	1782	1296	486

Table 9 provides some summary statistics for the entire sample and for the subsamples of immediate college entrants and delayed college entrants. Immediate and delayed college entrants differ in several dimensions. Immediate entrants are more likely to be female, have higher abstract and manual ASVAB subscores, and choose relatively

Table 10: Joint Skill Distribution: Abstract and Routine ASVAB Quartile Pairs

		Routine ASVAB Quartile			
		1	2	3	4
Abstract ASVAB Quartile	1	13.4%	5.9%	4.0%	1.7%
	2	6.0%	8.2%	6.5%	4.4%
	3	3.6%	6.6%	7.2%	7.6%
	4	2.1%	4.4%	7.4%	11.1%

abstract-intensive early-career occupations. Immediate entrants have higher family income and are more likely to have at least one parent who attended college. Finally, two noncognitive scores available in the NLSY79 are compared across individuals. The difference in the Rotter Score suggests delayed college entrants have, on average, a more external locus of control. The difference in Rosenberg score suggests that immediate college entrants have, on average, lower self-esteem.

Table 10 investigates the joint distribution of ASVAB subscores in the sample and displays the fraction of individuals in each of the sixteen abstract-routine skill quartile pairs. It is clear that abstract and routine scores are somewhat positively correlated. However, a significant number of individuals fit in each subscore quartile pair. This chapter explores how these abstract and routine skills separately affect future income trajectories, so this variation is crucial in the analyses below.

2.5 College Entry Timing, Early-Career Occupation Choices, and Future Income Trajectories

In order to better understand the dynamics of skill development through college and the early career, it is useful to isolate how these college entry timing choices and early-career occupation choices are associated with variation in observed income trajectories. To do so, a linear specification of log income w_{it} for individual i in year t is estimated. The full specification with estimates provided in Column (3) of Table 11 takes the following form.

$$w_{it} = \alpha_0 + \alpha_D D_i + \sum_{s=A,R} (\beta_s H_i^s + \gamma_{sD_i} T_i^s) + \Gamma X_{it} + \epsilon_{it} \tag{2.1}$$

This specification has the flexibility to capture differences arising from dynamic skill development through college and the early career. Skill development beyond college and the early career is modeled as a quadratic function of labor force experience included in a vector of additional covariates X_{it}. This specification is consistent with a large literature on this empirical specification of homogeneous human capital growth dating back to Mincer (1958).

Log income also depends upon delayed college entry D_i, abstract and routine skill measures at the time of high school graduation (H_i^A, H_i^R), abstract and routine task completion in the early career (T_i^A, T_i^R), and other controls included in the vector X_{it}. In the results presented, X_{it} includes interactions between sex and race along with the quadratic experience term. An unexplained portion of log income is modeled as an idiosyncratic shock ϵ_{it} assumed to be independent across individuals i. Standard errors are clustered at the individual level to correct for possible correlation in ϵ_{it} across time.

Columns (1) and (2) of Table 11 reveal that there is a significant future income penalty associated with delayed college entry. Both abstract and routine test scores positively affect future outcomes, as do abstract and routine task completion in the early career. Since abstract and routine test scores are normalized to have the same mean and standard deviation, comparing the coefficients can inform the relative importance of each of these factors. The results suggest that the abstract test score generally has a larger effect on productivity than routine test scores. Similarly, abstract task completion in the early career generally has a larger effect on productivity routine task completion.

The full specification also allows for one of the main flexibilities of the model explored in Section 2. Namely, the skill development associated with abstract and routine task completion in the early career may depend upon the timing of college entry. To capture this dynamic effect, interaction terms between task completion measures and a delayed college entry indicator enter into the model.

Column (3) of Table 11 provides estimates from a model that includes these interaction terms. As compared to immediate college entrants, delayed college entrants have significantly lower returns to abstract task completion. In contrast, delayed college entrants have higher returns to routine task completion, although the estimate of the interaction coefficient on routine task completion is not statistically significant. These results are informative about the dynamic skill development process through college and the early career. The results from Column (3) of Table 11 suggest that the abstract tasks preformed in the early career are most productive when performed *after* college as would an immediate college entrant, while the routine tasks performed in the early career are most productive when performed *before* college as would a delayed college entrant.

Table 11: Skill Determinants of Future Income Trajectories

	(1)	(2)	(3)
Delay	-0.223	-0.227	-0.221
	(0.0155)	(0.0155)	(0.0399)
Abstract ASVAB	0.143	0.104	0.102
	(0.0101)	(0.0104)	(0.0105)
Routine ASVAB	0.0848	0.0704	0.0721
	(0.00967)	(0.00988)	(0.00995)
Abstract Tasks		0.553	0.581
		(0.0239)	(0.0266)
Routine Tasks		0.170	0.147
		(0.0254)	(0.0287)
Abstract Tasks * Delay			-0.143
			(0.0586)
Routine Tasks * Delay			0.102
			(0.0567)
Observations	21143	21143	21143
R-squared	0.216	0.235	0.241

Standard errors in parentheses. Dependent variable is the log income for for college graduates with at least 10 years of labor force experience. Each specification also controls for quadratic experience and interactions for race, sex, and year of observation. Standard errors are clustered at the individual level.

Limitations

It is important to discuss several limitations in the interpretation of the estimates in Table 11. It may be the case that certain factors affecting earnings differences between immediate and delayed entrants are unobserved to the econometrician. This can introduce selection bias in the estimated coefficients, a serious concern for researchers attempting to estimate causal relationships using observational data. In this setting, there are multiple potential selection issues that warrant further discussion.

First, let us consider the problem of selection into college entry timing. Several factors could introduce this type of selection bias. For one, only the ASVAB subscores are used to measure skills at the time of high school graduation. Omission of additional variables that reflect differences in skills at the time of high school graduation can cause bias in the estimates. Another issue arises when college quality affects future earnings. The estimates do not account for variation in returns to college quality, but there is evidence that the returns to college depend upon the quality of the college attended.[4]

Next, there may also be selection into college degree completion. However, it is not clear that this type of selection causes an increase in bias. In particular, consider the case where delayed college entrants (graduates only) and immediate college entrants (graduates only) are more similar than all delayed entrants (graduates and non-graduates) and all immediate entrants (graduates and non-graduates). In this case, limiting the sample to college graduates would likely cause a reduction of the total effect of selection bias on the sample.

These selection problems are also acknowledged by Lin and Liu (2019). They use a

[4]The literature on returns to college quality is relatively large. A few examples include Black and Smith (2006) and Dale and Krueger (2013).

propensity score matching approach to estimate the treatment effect of delaying college entry on future earnings, and they estimate a negative and significant effect. To deal with selection on unobservables, they perform an analysis developed by Oster (2017) to explore the sensitivity of the estimate to selection on unobservables when it is proportional to selection on observables. Results suggest that the effect of delayed college entry is robust to a relatively large amount of selection on unobservables of this type.

For the results in Table 11, it is also important to recognize the potential impact of sorting into different early-career occupations on characteristics unobserved to the econometrician. For example, workers that are unobservably higher skill may sort into more abstract-intensive careers. Occupational sorting of this type could cause the estimates of Table 11 to overstate the effect of early-career occupational task completion on future income trajectories.

2.6 How Do Early-Career Occupation Choices Differ Across College Entry Timing?

Despite these limitations, the regression results in Table 11 illuminate some important ways in which expectations of future earnings trajectories can affect early-career occupation choices. Since the returns to abstract task completion are higher for immediate college entrants, the early-career occupation choices for immediate college entrants should be more abstract-intensive. Similarly, since the returns to routine task completion are higher for delayed college entrants, the early-career occupation choices for delayed college entrants should be more routine-intensive. The basic model also predicts that for either college entry timing choice, individuals seeking the highest returns to early-career task completion should invest more in abstract (routine) tasks when they have higher initial abstract (routine) skills at the time of high school graduation.

To test these basic predictions, it is beneficial to consider how individual early-career task content choices vary across college entry timing choices and abstract and routine skill endowments. Table 12 displays estimates from a regression of different early-career occupational task content measures on college entry timing decisions and abstract and routine test scores.

Column (1) of Table 12 estimates differences in abstract task content by college entry timing and test scores. Individuals with higher abstract test scores choose early-career occupations with higher abstract task content. This is true for both immediate and delayed college entrants. The link between delayed entry and abstract task content is negative but insignificant after controlling for test score measures. Column (2) of Table 12 estimates differences in routine task content by college entry timing and test scores. As predicted, the delayed entrants choose early-career occupations with significantly higher routine task content. Further, individuals with higher routine test scores are significantly more likely to choose higher routine task content in the early career.

Column (3) of Table 12 considers differences in a measure of the relative abstract intensity in the early-career occupation. This measure is simply constructed by taking the log of the difference between abstract and routine task content measured for each individual's early career occupation. As predicted, early-career delayed college entrants choose early-career occupations with significantly lower relative abstract intensity when compared to immediate college entrants. Relative abstract intensity is significantly higher for individuals with higher abstract test scores. Relative abstract intensity is lower for

Table 12: Task Intensities by Enrollment Timing and Skill Endowments

	(1) Abstract	(2) Routine	(3) Log Difference
Delay	-0.0425	0.416	-0.152
	(0.143)	(0.149)	(0.0664)
Abstract ASVAB	0.300	-0.0421	0.0904
	(0.0814)	(0.0853)	(0.0382)
Routine ASVAB	0.0993	0.238	-0.0199
	(0.0851)	(0.0868)	(0.0400)
Constant	3.066	3.658	-0.248
	(0.0825)	(0.0878)	(0.0397)
Observations	1782	1782	1782
R-squared	0.075	0.070	0.092

Standard errors in parentheses. Dependent variables are measures of abstract, routine, and relative early-career occupational task intensity.

individuals with higher routine test scores, but this association is not significant.

The results from Table 12 are largely consistent with the simple predictions of the model. Individuals with higher abstract (routine) skill endowments tend to invest more intensely in abstract (routine) tasks. Further, delayed college entrants tend to choose relatively less abstract-intensive and more routine-intensive early-career occupations than observably similar immediate college entrants.

2.7 Determinants of College Entry Timing Decisions

The goal of this final section is to illuminate the importance of different potential determinants of delayed college entry. Table 13 displays estimates from a logistic regression that explores how both abstract and manual skill measures at the time of high school graduation are associated with college entry delay. As discussed earlier in this chapter, several other factors are likely to contribute to the decision to delay college entry. For this reason, other determinants such as race, sex, family income, noncognitive assessments, and whether at least one parent attended college are also considered.

The results demonstrate that individuals with higher abstract test scores are significantly more likely to enter college immediately. In contrast, routine test scores are not significantly associated immediate college entry decisions. These patterns are robust to additional controls for sex, race, family income measures, and an indicator for whether or not either parent went to college. Results also suggest that skill measures at the time of high school graduation are not the only important predictors of delayed college entry. Race, family income, and an indicator for parental college enrollment are all significantly associated with immediate college entry.

The framework of dynamic skill development outlined in this chapter provides a simple interpretation of these results. In particular, the model demonstrates how differences in skills at the time of high school graduation can create tangible differences in the penalty to delayed college entry relative to immediate college entry. The log income regression estimates from Table 11 reveal that abstract task completion is most beneficial for skill development for immediate college entrants. The model therefore predicts that individuals with low abstract test scores have the least to lose from delaying college entry and accruing work experience. For this reason, other factors are more likely to affect the decision to delay college entry. In contrast, those with high abstract test scores face a significant penalty to entering the workforce after high school graduation rather than entering college immediately. For these individuals, the delay penalty is more likely to outweigh other factors that contribute to college entry timing decisions. These basic

Table 13: Determinants of College Entry Delay: Logistic Regression

	(1)	(2)
Abstract ASVAB	-0.599	-0.559
	(0.0809)	(0.105)
Routine ASVAB	0.0664	-0.00283
	(0.0887)	(0.0961)
Male		0.0154
		(0.132)
Black		0.725
		(0.169)
Family Income/1000		-0.0194
		(0.00449)
One Parent College		-0.617
		(0.144)
Rotter Score		0.0372
		(0.0285)
Rosenberg Score		-0.0219
		(0.0165)
Constant	-0.988	-0.579
	(0.0707)	(0.506)
Observations	1739	1634

Standard errors in parentheses.

insights can provide some direction for policymakers who seek to minimize the negative future income effects of delayed college entry. In particular, policies specifically targeted at encouraging immediate college entry of high school graduates with high abstract test scores but other characteristics that put them at higher risk for delayed college enrollment are most likely to be effective in reducing the negative effects of delayed college entry.

This recommendation is consistent with the success of several policy interventions targeting high-achieving but low-income or minority students. For example, Hyman (2018) examines a study in which nearly 50,000 high school seniors in which treated individuals were mailed a letter encouraging them to apply to college and providing the web address that provides information about college enrollment. The study showed that very high-achieving poor and minority students were most likely to browse the website and had the largest increase in probability of enrolling in college from treatment. Dynarski et al (2018) study the HAIL Scholarship that targets high-achieving, low-income high school seniors and provides a select group with the promise of tuition subsidy for attending the University of Michigan. Of the documented 15 percent increase in enrollment effect, 8 percent consisted of high school graduates who would not have otherwise enrolled in any four-year college. The success of these incentives in inducing immediate college entry for high-achieving students who face other barriers to college entry may stem from the particularly large gains associated with immediate college entry for these students.

2.8 Conclusion

This chapter explores the life-cycle earnings implications of delaying college entry through the lens of dynamic skill development. Delayed college entry can affect future earnings when early-career occupational investments completed before and after college affect skill development differently. On average, high school graduates who delay college entry tend to have lower earnings trajectories than observably similar immediate college entrants. Abstract-intensive early-career occupations are most beneficial for skill development after college, while routine-intensive early-career occupations are most beneficial for skill development before college. Accordingly, delayed college entrants choose relatively routine-intensive occupations before college entry, and immediate college entrants choose relatively abstract-intensive occupations after college completion. Individuals with high abstract test scores at the time of high school graduation face the largest penalty to delayed college entry. These results suggest that high school graduates respond to incentives that arise from dynamic skill development when choosing early-career occupations and the timing of college entry. This can inform policymakers who seek to minimize the negative future income effects of delayed college entry.

Bibliography

[1] Arcidiacono, P., Aucejo, E., Maurel, A. P., and Ransom, T. (2016). College Attrition and the Dynamics of Information Revelation. Working Paper No. 222. Economic Research Initiatives at Duke (ERID).

[2] Autor, D. and Dorn, D. (2013). The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review*, 103(5):1553-1597.

[3] Black, D. A. and Smith, J. A. (2006). Estimating the Returns to College Quality with Multiple Proxies for Quality. *Journal of Labor Economics*, 24(3):701-728.

[4] Bozick, R. and DeLuca, S. (2005). Better Late Than Never? Delayed Enrollment in the High School to College Transition. *Social Forces*, 84(1):531-554.

[5] Cameron, S. V. and Taber, C. (2004). Estimation of Educational Borrowing Constraints Using Returns to Schooling. *Journal of Political Economy*, 112(1):132-182.

[6] Carneiro, P. and Heckman, J. J. (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *The Economic Journal*, 112(482):989-1018.

[7] Dale, S. B. and Krueger, A. B. (2014). Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data. *Journal of Human Resources*, 49(2):323-358.

[8] Dynarski, S., Libassi C. J., Michelmore, K., and Owen, S. (2018). Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students. NBER Working Paper No. 25349.

[9] Griliches, Z. (1980). Schooling Interruption, Work While in School and the Returns from Schooling. *The Scandinavian Journal of Economics*, 82(2):291-303.

[10] Goldrick-Rab, S. (2006). Following their every move: An investigation of social-class differences in college pathways. *Sociology of Education*, 79(1):67-79.

[11] Goldrick-Rab, S. and Han, S. W. (2011). Accounting for socioeconomic differences in delaying the transition to college. *The Review of Higher Education*, 34(3):423-445.

[12] Heckman, J. J. and Robb, R. (1986). Alternative Methods for Evaluating the Impact of Interventions: An Overview. *Journal of Econometrics*, 30:239-267.

[13] Horn, L., Cataldi, E. F., and Sikora, A. (2005). Waiting to attend college: Undergraduates who delay their post-secondary enrollment (NCES 2005-152). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

[14] Hyman, J. (2018). Nudges, College Enrollment, and College Persistence: Evidence From a Statewide Experiment in Michigan. Working Paper 2018-10, University of Connecticut, Department of Economics.

[15] James, J. (2012). Delaying Enrollment and College Completion. *Federal Reserve Bank of Cleveland*.

[16] Johnson, M. T. (2013). Borrowing Constraints, College Enrollment, and Delayed Entry. *Journal of Labor Economics*, 31(4):669-725.

[17] Kane, T. J. (1996). College cost, borrowing constraints and the timing of college entry. *Eastern Economic Journal*, 22(2):181-194.

[18] Light, A. (1995). The effects of interrupted schooling on wages. *Journal of Human Resources*, 30(3):472-502.

[19] Lin, Y. and Liu, V. Y. T. (2019). Timing Matters: How Delaying College Enrollment Affects Earnings Trajectories. CCRC Working Paper No. 105, Teachers College, Columbia University.

[20] Lise, J. and Postel-Vinay, F. (2018). Multidimensional Skills, Sorting, and Human Capital Accumulation. Working Paper No. 131, Heller-Hurwicz Economics Institute, University of Minnesota.

[21] Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66(4):281-302.

[22] Oster, E. (2017). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2):187-204.

[23] Pugatch, T. (2018). Bumpy Rides: School-To-Work Transitions in South Africa. *LABOUR*, 32(2):205-242.

[24] Sanders, C. (2016). Skill Accumulation, Skill Uncertainty, and Occupational Choice. Working Paper.

[25] Wine, J., Bryan, M., and Siegel, P. (2013). 2011-12 National Postsecondary Student Aid Study (NPSAS:12): Data file documentation (NCES 2014-182). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

[26] Yamaguchi, S. (2012). Tasks and Heterogeneous Human Capital. *Journal of Labor Economics*, 30(1):1-53.

Appendix A

Technical Appendix

Part 1: Incorporating Race and Education-Specific Signaling Technologies

This part re-evaluates the prediction of the basic model in the case that the signal distribution varies by race or and education group. Let $u \sim \mathcal{N}(0, \sigma_u^2(r,s))$ denote the signal noise distribution that firms receive for workers with race r and schooling level s. As in the basic model presented above, firms observe $I_1 = \{r, s, \eta, \tilde{q}, \bar{a}_r(s)\}$ and offer a log wage schedule equal to the log of expected productivity.

$$log(w) = g(s) + \lambda(r,s)\tilde{q} + (1 - \lambda(r,s))(\bar{a}_r(s) + \frac{1}{2}\sigma_\epsilon^2) + \eta$$

where

$$\lambda(r,s) = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_u^2(r,s)}$$

The conditional expected log wage given econometrician observables $I_2 = \{r, s, a, \bar{a}_r(s)\}$ can then be derived.

$$E[log(w)|I_2] = g(s) + \lambda(r,s)a + (1 - \lambda(r,s))(\bar{a}_r(s) + \frac{1}{2}\sigma_\epsilon^2) + E[\eta|r,s]$$

The racial difference in measured returns to home inputs can be expressed as

$$\frac{\partial E[log(w)|a,s,W]}{\partial a} - \frac{\partial E[log(w)|a,s,B]}{\partial a} = \lambda(W,s) - \lambda(B,s)$$

A straightforward proposition about racial differences in returns to home environment is now derived.

Proposition A1: *If the signal distributions do not vary by race, meaning that* $\sigma_u^2(B,s) = \sigma_u^2(W,s)$ *for all* s*, then the returns to home inputs does not vary by race. Further, if firms have a clearer signal of productivity for a given racial group, then the returns to home inputs are higher for that racial group.*

The data provide strong evidence that the returns to the HOME score are not different across racial groups. In this context, race-independent productivity signal distributions are a necessary and sufficient condition for equal returns to home inputs, so Proposition A1 suggests that $\sigma_u^2(B,s) = \sigma_u^2(W,s)$ for all s. For this reason, we make the guided assumption signal distributions are race-independent and with noise variance $\sigma_u^2(s)$ and ratio $\lambda(s)$ for all s.

Racial differences in the returns to education follow

$$\frac{\partial E[log(w)|a,s,B]}{\partial s} - \frac{\partial E[log(w)|a,s,W]}{\partial s} =$$
$$\lambda'(s)(\bar{a}_W(s) - \bar{a}_B(s)) + (1-\lambda(s))(\bar{a}'_B(s) - \bar{a}'_W(s)) + (\frac{\partial E[\eta|s,W]}{\partial s} - \frac{\partial E[\eta|s,B]}{\partial s})$$

There are two contributions to racial differences in the returns to education. Both contributions stem from racial differences in the signaling value of education. First, consider a situation in which $\sigma_u^2(s)$ is decreasing in s, implying that $\lambda'(s) > 0$. As workers attain more education, the firm relies less on the prior belief $\bar{a}_r(s)$ and more on the signal $\tilde{q}$. If the firm's prior is lower for black workers than for white workers at each education level, then increasing education has more value for black workers than on white workers (even when the signaling technologies do not differ by race). This is the driving mechanism of Arcidiacono et al (2010).

The second contribution is the same as the one in the baseline model, where we assume that $\bar{a}'_B(s) > \bar{a}'_W(s)$. The third contribution due to differences in η is also the same as in the baseline model.

The principal model prediction is summarized in the following proposition.

Proposition A2: *Let $S = [s_L, s_H]$ be a range of education levels over which*

(i) $\sigma_u^2(B, s) = \sigma_u^2(W, s) = \sigma_u^2(s)$

(ii) $\sigma_u^2(s)$ is *non-increasing* in s

(iii) $\bar{a}_B(s) < \bar{a}_W(s)$

(iv) $\bar{a}'_B(s) > \bar{a}'_W(s)$

(v) $\frac{\partial E[\eta|s,W]}{\partial s} \geq \frac{\partial E[\eta|s,B]}{\partial s}$

Then, if $\frac{\partial E[log(w)|a,s,B]}{\partial s} > \frac{\partial E[log(w)|a,s,W]}{\partial s}$ for all a and all $s \in S$, it follows that $1 - \lambda > 0$ and the higher measured wage returns to educational attainment for black workers over the range S are driven by statistical discrimination.

Assumption (i) implies race-independent signaling technologies. Assumptions (iii) and (iv) are consistent with racial gaps in home inputs across educational attainment. Assumption (v) is identical to the assumption made in the baseline model and explored in this paper.

Assumption (ii) implies that the signaling technology is *not* noisier at higher levels of educational attainment. As mentioned, this is consistent with the results in Arcidiacono et al (2010). The intuition provided is that for higher education workers, employers are more likely to receive more detailed information on grades, specific past experiences, skills acquired, and test scores. This is likely to result in a non-increasing $\sigma_u^2(s)$.

Part 2: Incorporating Racial Bias in Home Input Measures

Here, a systematic racial bias in measurement is incorporated into the model. As in the baseline model, true productivity follows

$$log(p^{\star}) = g(s) + a + \eta + \epsilon$$

While productivity depends upon the true measure of home inputs a, the econometrician instead observes

$$\tilde{a} = a + b$$

where b represents some bias in the measure. The employers' signal of unobserved productivity is

$$\tilde{q} = a + \epsilon + u$$

The derivation of the racial difference in measured wage returns to education is similar to that presented in the baseline model. This difference is given by

$$\frac{\partial E[log(w)|\tilde{a}, s, B]}{\partial s} - \frac{\partial E[log(w)|\tilde{a}, s, W]}{\partial s} =$$
$$(1-\lambda)(\bar{a}'_B(s) - \bar{a}'_W(s)) + (\frac{\partial E[\eta|s, W]}{\partial s} - \frac{\partial E[\eta|s, B]}{\partial s}) + (\frac{\partial E[b|W, s]}{\partial s} - \frac{\partial E[b|B, s]}{\partial s})$$

On the right hand side, the first and second terms are identical to those in the baseline model, while the third term represents the effect of persistent racial bias in HOME score measurement. If both the second and third term are non-positive, racial differences in the returns to education must be a result of the first term, which represents statistical discrimination. The third term violates the assumption if and only if racial bias $E[b|W, s] - E[b|B, s]$ is smallest at the lowest levels of education and increases with educational attainment.

Appendix B

Supplementary Figures and Tables

Figure 4: Average HOME Scores by Race and Highest Grade

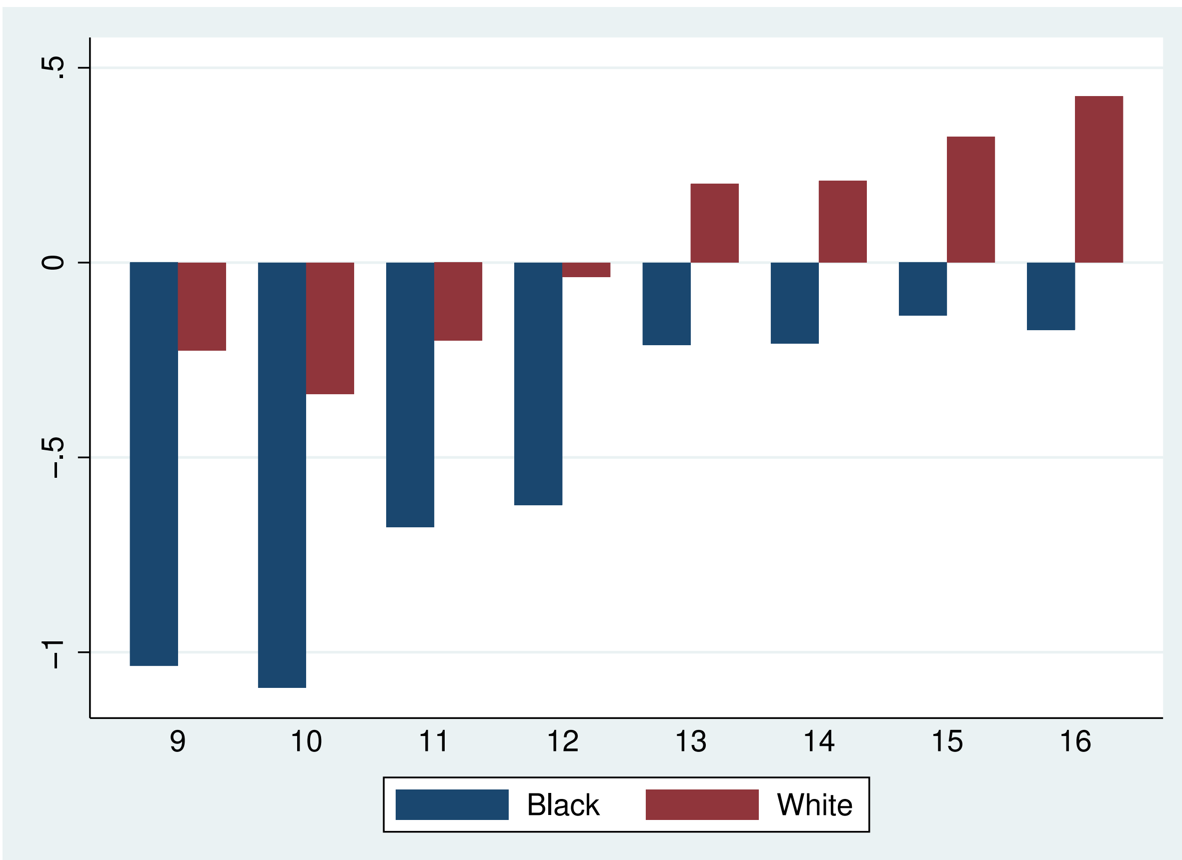
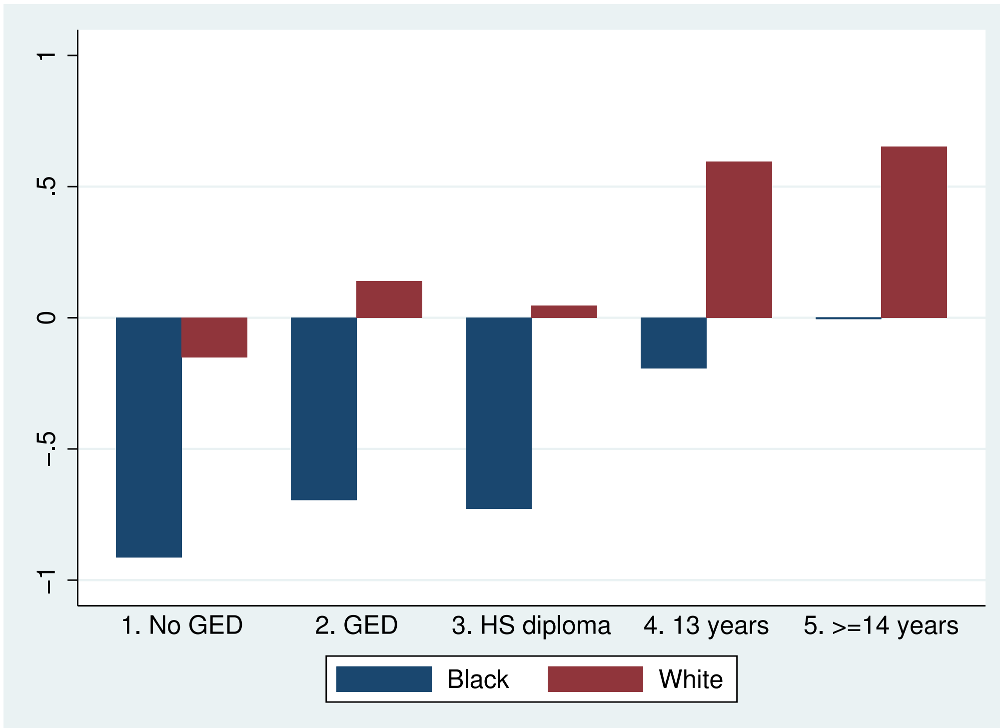


Figure 5: Average PIAT Scores by Race and Highest Grade

(a) Math PIAT Scores



(b) Reading PIAT Scores

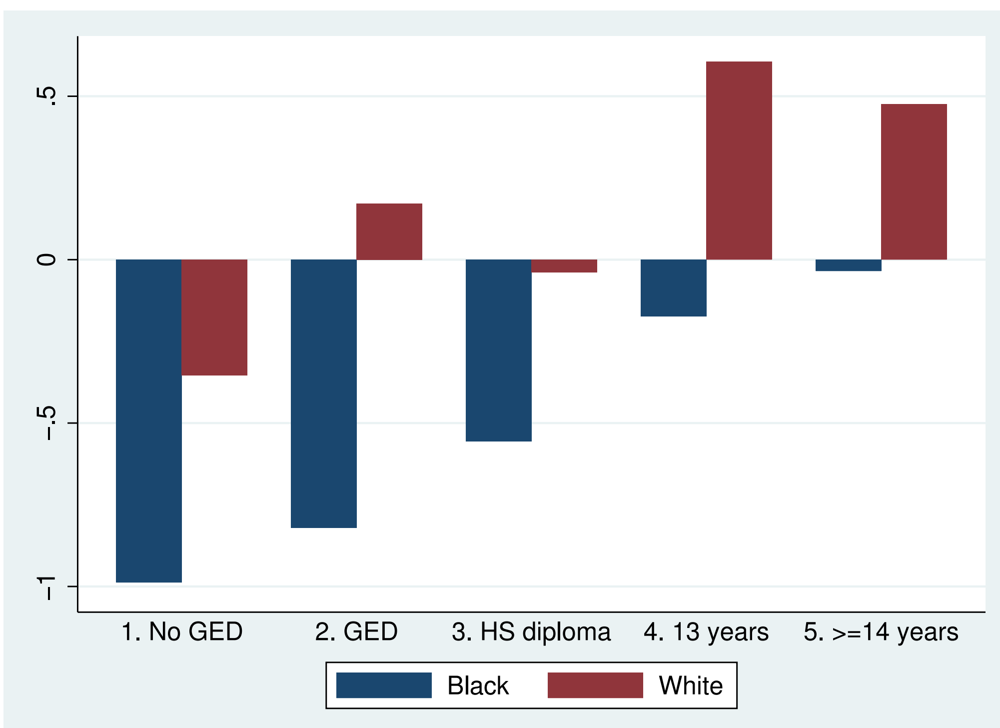
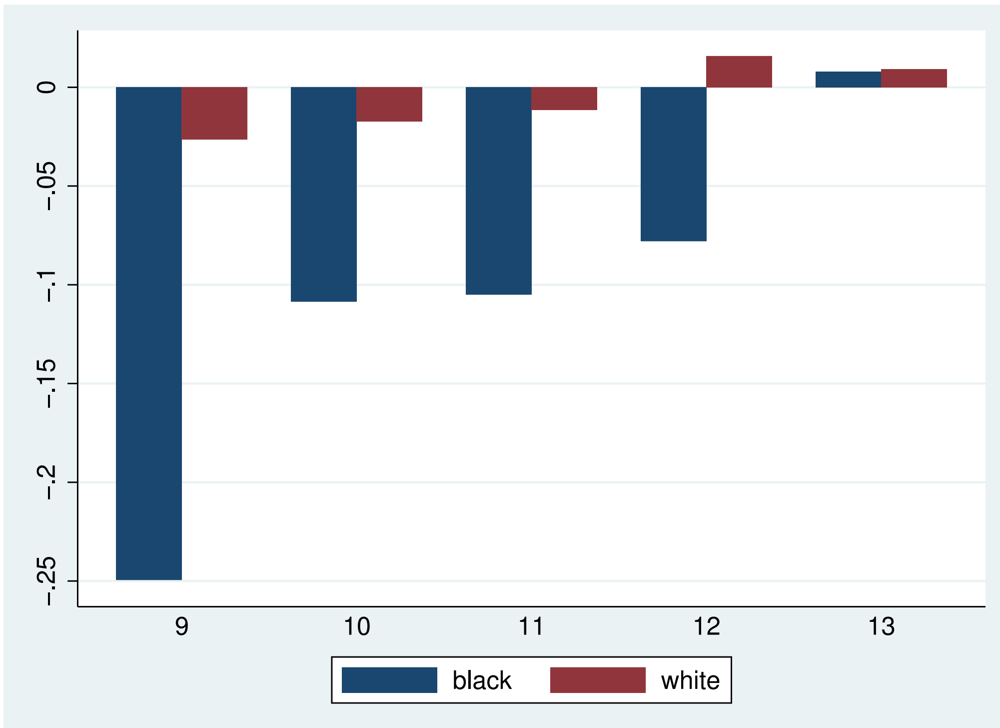


Figure 6: Average Adjusted Log Wages by Race and Highest Grade

(a) Adjusted for HOME



(b) Adjusted for HOME, PIAT, and Neighborhood

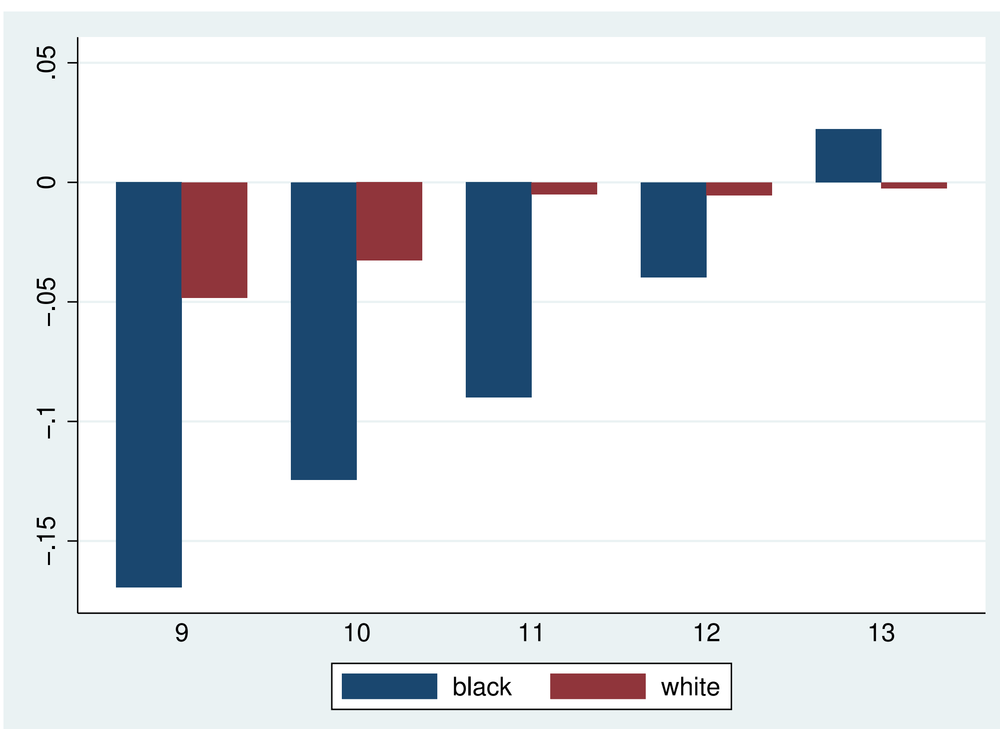


Table 14: Log Wage Regressions: Robustness to Alternative Controls

	(1)	(2)	(3)	(4)
Black	-0.0512	-0.0400	-0.0421	-0.0285
	(0.0376)	(0.0477)	(0.0394)	(0.0487)
HOME Score	0.0670	0.0679	0.0624	0.0626
	(0.0151)	(0.0178)	(0.0157)	(0.0183)
High Grade	0.0403	0.0361	0.0395	0.0357
	(0.0190)	(0.0234)	(0.0190)	(0.0233)
Grade*Black	0.0400	0.0471	0.0423	0.0488
	(0.0230)	(0.0265)	(0.0234)	(0.0268)
PIAT Math		0.0174		0.0165
		(0.0220)		(0.0215)
PIAT Reading		-0.00288		-0.00375
		(0.0144)		(0.0144)
Neighborhood (Rating for Kids)			0.0443	0.0508
			(0.0360)	(0.0396)
Neighborhood (Supervision)			-0.0233	-0.0344
			(0.0371)	(0.0416)
Neighborhood (Crime)			0.00717	0.0574
			(0.0470)	(0.0510)
Neighborhood (Neighbors Care)			0.00617	0.00453
			(0.0742)	(0.0868)
Observations	721	623	721	623
R-squared	0.133	0.144	0.136	0.150

Standard errors in parentheses. Dependent variable is log of hourly wages observed between ages 22 and 27. All specifications also control for an experience term and year at wage observation. Standard errors are clustered at the individual level. Only those workers with less than 14 years of education are included in the sample Neighborhood quality variables are survey answers of children's mothers from the NLSY79 data. PIAT scores are the oldest nonmissing observation for each child and age-adjusted. The coefficient estimates for Black are centered at 13 years of education.