LABOR MARKET DISCRIMINATION AND RACIAL DIFFERENCES IN PREMARKET FACTORS*

PEDRO CARNEIRO,
University College London

JAMES J. HECKMAN,
University of Chicago

and

DIMITRIY V. MASTEROV
University of Chicago

Abstract

This paper investigates the relative significance of differences in cognitive skills and discrimination in explaining racial/ethnic wage gaps. We show that cognitive test scores for exams taken prior to entering the labor market are influenced by schooling. Adjusting the scores for racial/ethnic differences in education at the time the test is taken reduces their role in accounting for the wage gaps. We also consider evidence on parental and child expectations about education and on stereotype threat effects. We find both factors to be implausible alternative explanations for the gaps we observe. We argue that policies need to address the sources of early skill gaps and to seek to influence the more malleable behavioral abilities in addition to their cognitive counterparts. Such policies are far more likely to be effective in promoting racial and ethnic equality for most groups than are additional civil rights and affirmative action policies targeted at the workplace.

I. Introduction

It is well documented that civil rights policy directed toward the South raised black economic status in the 1960s and 1970s.[1] Yet substantial gaps remain in the market wages of African-American males and females com-

* This research was supported by a grant from the American Bar Foundation and National Institutes of Health grant R01-HD043411. We thank an anonymous referee for helpful comments. Carneiro was supported by Fundação Ciência e Tecnologia and Fundação Calouste Gulbenkian. We thank Derek Neal and Rodrigo Soares for helpful comments and Maria Isabel Larenas, Maria Victoria Rodriguez, and Xing Zhong for excellent research assistance.

[1] John J. Donohue & James J. Heckman, Continuous versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks, 29 J. Econ. Literature 1603 (1991).

[*Journal of Law and Economics*, vol. XLVIII (April 2005)]
 0022-2186/2005/4801-0001$01.50

pared with those of white males and females.[2] There are sizable wage gaps for Hispanics as well.

Columns 1 of Table 1 report, for various ages, the mean hourly log wage gaps for a cohort of young black and Hispanic males and females.[3] The reported gaps are not adjusted for differences in schooling, ability, or other market productivity traits. The table shows that, on average, black males earned wages that were 25 percent lower than those of white males in 1990. Hispanic males earned wages that were 17.4 percent lower in the same year. The gaps increase for males as the cohort ages. For women, there are smaller gaps for blacks and virtually no gap at all for Hispanics, and the gaps for women show no clear trend with age.[4] Joseph Altonji and Rebecca Blank[5] report similar patterns using data from the March supplement to the Current Population Survey for 1968–98.

These gaps are consistent with claims of pervasive labor market discrimination against minorities. Minority workers with the same ability and training as white workers may be receiving lower wages. There is, however, another equally plausible explanation consistent with the same evidence. Minorities may bring less skill and ability to the market. Although there may be discrimination or disparity in the development of these valuable skills, the skills may be rewarded equally across all demographic groups in the labor market. Clearly, a variety of intermediate explanations that combine both hypotheses are consistent with the data just presented.

The two polar interpretations of market wage gaps have profoundly different policy implications. If persons of identical skill are treated differently in the labor market on the basis of race or ethnicity, a more vigorous enforcement of civil rights and affirmative action in the marketplace would appear to be warranted. On the other hand, if the gaps are solely due to unmeasured abilities and skills that people bring to the labor market, then a redirection of policy toward fostering skills should be emphasized as opposed to a policy of ferreting out discrimination in the workplace.

[2] The literature on African-American economic progress in the twentieth century is surveyed in James J. Heckman & Petra Todd, Understanding the Contribution of Legislation, Social Activism, Markets and Choice to the Economic Progress of African Americans in the Twentieth Century (unpublished manuscript, Am. Bar Found. 2001).

[3] These gaps are for a cohort of young persons aged 26–28 in 1990 from the National Longitudinal Survey of Youth of 1979 (NLSY79). They are followed for 10 years until they reach ages 36–38 in 2000.

[4] However, the magnitudes (but not the direction) of the female gaps are less reliably determined, at least for black women. Derek Neal, The Measured Black-White Wage Gap among Women Is Too Small, 112 J. Pol. Econ. S1 (2004), shows that racial wage gaps for black women are underestimated by these types of regressions since they do not control for selective labor force participation. This same line of reasoning is likely to hold for Hispanic women.

[5] Joseph Altonji & Rebecca Blank, Gender and Race in the Labor Markets, in 3C Handbook of Labor Economics 3143 (Orley Ashenfelter & David Card eds. 1999).

Derek Neal and William Johnson[6] shed light on the relative empirical importance of market discrimination and skill disparity in accounting for wage gaps by race. Controlling for scholastic ability measured in the mid-teenage years, they substantially reduce but do not fully eliminate wage gaps for black males in 1990–91 data. They more than eliminate the gaps for black females. Columns 2 in Table 1 show our version of the estimates reported in the Neal-Johnson study, expanded to cover additional years.[7] For black males, controlling for an early measure of ability cuts the black-white wage gap in 1990 by 76 percent. For Hispanic males, controlling for ability essentially eliminates the wage gap with whites. For women, the results are even more striking. Wage gaps are actually reversed, and controlling for ability produces higher wages for minority females. This evidence suggests that the endowments people bring to the labor market play a substantial role in accounting for minority wage gaps.

This paper critically examines the Neal-Johnson argument and brings fresh evidence to bear on it. With some important qualifications, our analysis supports their conclusion that factors determined outside of the market play the major role in accounting for minority-majority wage differentials in modern labor markets.

In producing the wage gaps shown in Table 1, we follow a practice suggested by Neal and Johnson and do not adjust for the effects of racial and economic differences in schooling, occupational choice, or work experience on wages. Racial and ethnic differences in these factors may reflect responses to labor market discrimination and should not be controlled for in regressions that estimate the "full effect" of race on wages through all channels since doing so may spuriously reduce estimated wage gaps by introducing a proxy for discrimination into the control variables. While the motivation for their procedure is clear, their qualitative claim is false. Including schooling in a wage regression raises the estimated wage gaps and produces more evidence of racial disparity. Gaps when schooling is fixed and not fixed are both of interest and answer different questions.

Gaps in measured ability by ethnicity and race are substantial. Figure 1 plots the ability distribution as measured by age-corrected Armed Forces Qualification Test (AFQT) scores[8] for males and females in the National

[6] Derek Neal & William Johnson, The Role of Premarket Factors in Black-White Wage Differences, 104 J. Pol. Econ. 869 (1996).

[7] We use a sample very similar to the one used in their study. It includes individuals born only in 1962–64. This exclusion is designed to alleviate the effects of differential schooling at the test date on test performance and to ensure that the AFQT is taken before the individuals enter the labor market, so that it is more likely to be a premarket factor.

[8] Age-corrected AFQT is the standardized residual from the regression of the AFQT score on age at the time of the test dummy variables. AFQT is a subset of four out of 10 Armed Services Vocational Aptitude Battery (ASVAB) tests used by the military for enlistment screening and job assignment. It is the summed score from the word knowledge, paragraph comprehension, mathematics knowledge, and arithmetic reasoning ASVAB tests.

TABLE 1

CHANGE IN THE BLACK-WHITE LOG WAGE GAP INDUCED BY CONTROLLING FOR AGE-CORRECTED AFQT SCORES, 1990–2000

	1990		1991		1992		1993		1994		1996		1998		2000	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Males:																
Black	−.250	−.060	−.251	−.082	−.302	−.113	−.282	−.104	−.286	−.093	−.373	−.149	−.333	−.069	−.325	−.089
	(.028)	(.030)	(.028)	(.030)	(.029)	(.030)	(.028)	(.030)	(.031)	(.033)	(.032)	(.034)	(.034)	(.035)	(.035)	(.035)
Hispanic	−.174	−.035	−.113	.020	−.146	−.014	−.159	−.027	−.143	.005	−.186	−.031	−.195	−.006	−.215	−.053
	(.032)	(.032)	(.032)	(.033)	(.033)	(.032)	(.032)	(.032)	(.036)	(.036)	(.038)	(.038)	(.040)	(.038)	(.040)	(.038)
Age	. . .	.050	. . .	.030	. . .	.038	. . .	.030	. . .	.023	. . .	.020	. . .	.014	. . .	.008
	. . .	(.014)	. . .	(.014)	. . .	(.014)	. . .	(.014)	. . .	(.016)	. . .	(.016)	. . .	(.017)	. . .	(.017)
AFQT	. . .	.183	. . .	.161	. . .	.179	. . .	.172	. . .	.188	. . .	.216	. . .	.254	. . .	.241
	. . .	(.013)	. . .	(.013)	. . .	(.013)	. . .	(.013)	. . .	(.014)	. . .	(.015)	. . .	(.015)	. . .	(.015)
$AFQT^2$	. . .	−.022	. . .	−.007	. . .	−.001	. . .	.002	. . .	.014	. . .	.021	. . .	.032	. . .	.023
	. . .	(.011)	. . .	(.011)	. . .	(.011)	. . .	(.011)	. . .	(.012)	. . .	(.013)	. . .	(.013)	. . .	(.013)
Intercept	2.375	.957	2.372	1.463	2.404	1.202	2.423	1.431	2.458	1.652	2.533	1.756	2.589	1.960	2.629	2.224
	(.017)	(.385)	(.017)	(.406)	(.017)	(.422)	(.017)	(.432)	(.018)	(.488)	(.019)	(.540)	(.020)	(.594)	(.020)	(.621)
N	1,538	1,505	1,553	1,514	1,536	1,503	1,542	1,504	1,522	1,485	1,554	1,519	1,494	1,462	1,438	1,404

Females:																
Black	−.172	.041	−.200	.030	−.201	.010	−.167	.093	−.148	.099	−.147	.132	−.201	.071	−.200	.069
	(.031)	(.033)	(.032)	(.035)	(.031)	(.033)	(.035)	(.037)	(.035)	(.037)	(.035)	(.036)	(.034)	(.036)	(.036)	(.038)
Hispanic	−.003	.154	−.017	.153	−.059	.114	.009	.198	−.018	.170	−.006	.193	−.069	.151	−.064	.149
	(.035)	(.035)	(.037)	(.038)	(.036)	(.035)	(.039)	(.039)	(.040)	(.040)	(.041)	(.039)	(.039)	(.039)	(.041)	(.041)
Age	. . .	.010	. . .	.038	. . .	.016	. . .	.016	. . .	.008	. . .	−.009	. . .	.013	. . .	−.018
	. . .	(.015)	. . .	(.017)	. . .	(.016)	. . .	(.017)	. . .	(.018)	. . .	(.018)	. . .	(.017)	. . .	(.018)
AFQT	. . .	.217	. . .	.234	. . .	.229	. . .	.271	. . .	.267	. . .	.283	. . .	.274	. . .	.273
	. . .	(.016)	. . .	(.018)	. . .	(.017)	. . .	(.018)	. . .	(.019)	. . .	(.018)	. . .	(.018)	. . .	(.018)
$AFQT^2$	. . .	.005	. . .	.000	. . .	−.001	. . .	−.012	. . .	−.024	. . .	.005	. . .	.000	. . .	−.008
	. . .	(.014)	. . .	(.014)	. . .	(.014)	. . .	(.015)	. . .	(.016)	. . .	(.015)	. . .	(.015)	. . .	(.015)
Intercept	2.141	1.750	2.175	.982	2.193	1.615	2.174	1.558	2.218	1.858	2.246	2.412	2.311	1.724	2.339	2.867
	(.019)	(.420)	(.020)	(.467)	(.019)	(.458)	(.021)	(.520)	(.022)	(.555)	(.022)	(.582)	(.021)	(.603)	(.022)	(.663)
N	1,356	1,325	1,335	1,299	1,317	1,278	1,319	1,281	1,318	1,288	1,381	1,344	1,370	1,329	1,316	1,276

NOTE.—This table reports coefficients of regressions of log hourly wages in each year on the variables listed in the left-hand column. Values given are for participants in the National Longitudinal Survey of Youth born after 1961. Age-corrected Armed Forces Qualification Test (AFQT) is the standardized residual from the regression of the AFQT score on age at the time of the test dummy variables. AFQT is a subset of four out of 10 Armed Services Vocational Aptitude Battery (ASVAB) tests used by the military for enlistment screening and job assignment. It is the summed score from the word knowledge, paragraph comprehension, mathematics knowledge, and arithmetic reasoning ASVAB tests. All wages are in 1993 dollars. The coefficients on the AFQT variables represent the effect of a 1-standard-deviation increase in the score on the log hourly wage. Since the wage is measured in log points, the gaps for blacks and Hispanics correspond approximately to percentage point differences relative to the white mean; that is, the black-white gap of −.25 in 1990 corresponds to wages for blacks in that year that are 25% lower than those for whites.

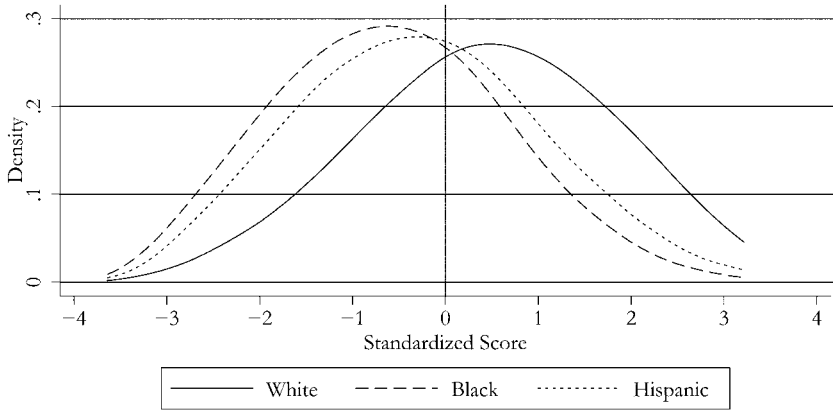

FIGURE 1.—Density of age-corrected AFQT scores for NLSY79 males born after 1961

Longitudinal Survey of Youth of 1979 (NLSY79).[9] As noted by Richard Herrnstein and Charles Murray,[10] ability gaps are a major factor in accounting for a variety of racial and ethnic disparities in socioeconomic outcomes. Stephen Cameron and James Heckman[11] show that controlling for ability, blacks and Hispanics are more likely to enter college than are whites.[12]

Neal and Johnson[13] argue that ability measured in the teenage years is a "premarket" factor, meaning that it is not affected by expectations or actual experiences of discrimination in the labor market. They offer no explicit criterion for determining which factors are premarket and which are not.

Schooling affects test scores,[14] and levels of minority schooling are lower than white schooling levels, both generally and in the samples used by Neal and Johnson. Their test score is contaminated by schooling attainment at the

[9] In our Web appendix (http://jenni.uchicago.edu/JLE/FILES/JLE_Appendix_2004-07-20_dvm.pdf), we show that the same patterns emerge when we divide the sample by gender.

[10] Richard Herrnstein & Charles Murray, The Bell Curve (1994).

[11] Stephen V. Cameron & James J. Heckman, The Dynamics of Educational Attainment for Black, Hispanic, and White Males, 109 J. Pol. Econ. 455 (2001).

[12] Sergio Urzua, The Educational White-Black Gap: Evidence on Years of Schooling (Working paper, Univ. Chicago, Dep't Econ. 2003), shows that this effect arises from greater minority enrollment in 2-year colleges. Controlling for ability, whites are more likely to attend and graduate from 4-year colleges. Using the Current Population Survey, Sandra E. Black & Amir Sufi, Who Goes to College? Differential Enrollment by Race and Family Background (Working Paper No. w9310, Nat'l Bur. Econ. Res. 2002), finds that equating the family background of blacks and whites eliminates the black-white gap in schooling only at the bottom of the family background distribution. Furthermore, the gaps are eliminated in the 1980s but not in the 1990s.

[13] Neal & Johnson, *supra* note 6.

[14] See Karsten Hansen, James J. Heckman, & Kathleen Mullen, The Effect of Schooling and Ability on Achievement Test Scores, 121 J. Econometrics 39 (2004).

date of the test. When their test scores are adjusted for this factor, adjusted wage gaps increase.

The gaps in ability evident in Figure 1 stem in part from lower levels of schooling by minorities at the time of the test and may also arise from lowered academic effort in anticipation of future discrimination in the labor market. If skills are not rewarded fairly, the incentive to acquire them is diminished for those subject to prejudicial treatment. Discrimination in the labor market might sap the incentives of children and young adults to acquire skills and abilities but may also influence the efforts they exert in raising their own offspring. This means that even after adjusting their test scores for schooling, measured ability may not be a true premarket factor. Neal and Johnson[15] mention this qualification in their original paper, and their critics have subsequently reiterated it.

The gaps in ability may also be a consequence of adverse environments. Even if all wage gaps are due to ability, uncontaminated by expectations of market discrimination, the appropriate policy for eliminating ability gaps is not apparent from Table 1. Should policies focus on early ages through enriched Head Start programs or on improving schooling quality and reducing school dropout and repetition rates that plague minority children at later ages?

This paper demonstrates that ability gaps open up very early. Minorities enter school with substantially lower measured ability than whites. The black-white ability gap widens as the children get older and obtain more schooling, but the contribution of formal education to the widening of the gap is small when compared to the size of the initial gap. There is a much smaller widening of the Hispanic-white ability gap with schooling.

Our evidence and that of James Heckman, Maria Isabel Larenas, and Sergio Urzua[16] suggest that school-based policies are unlikely to have substantial effects on eliminating minority ability gaps. Factors that operate early in the life cycle of the child are likely to have the greatest impact on ability.

The early emergence of ability gaps indicates that child expectations play only a limited role in accounting for such gaps since very young children are unlikely to have formed expectations about labor market discrimination and to make decisions based on those expectations. However, parental expectations of future discrimination may still play a role in shaping children's outcomes.

The early emergence of measured ability differentials also casts doubt on

[15] Neal & Johnson, *supra* note 6.

[16] James Heckman, Maria Isabel Larenas, & Sergio Urzua, Accounting for the Effect of Schooling and Abilities in the Analysis of Racial and Ethnic Disparities in Achievement Test Scores (Working paper, Univ. Chicago, Dep't Econ. 2004).

the empirical importance of the "stereotype threat"[17] as a major factor contributing to black-white test score differentials. The literature on this topic claims that black college students at selective colleges perform worse on tests when they are told that the tests may be used to confirm stereotypes about black-white ability differentials. The empirical importance of this effect is in dispute in the psychology literature.[18]

The children in our data are tested at very young ages and are unlikely to be aware of stereotypes about minority inferiority or be affected by the stereotype threat that has been empirically established only for students at elite colleges. In addition, large gaps in test scores are also evident for Hispanics, a group for whom the stereotype threat has not been documented. The stereotype threat literature claims that measured test scores for minorities understate their true ability. Unless the effect is uniform across ability levels, incremental ability should be rewarded differently between blacks and whites. We find no evidence of such an effect.

Adjusting for the schooling attainment of minorities at the time that they take tests provides an empirically important qualification to the Neal-Johnson study.[19] An extra year of schooling has a greater impact on test scores for whites and Hispanics than for blacks. Adjusting the test score for schooling disparity at the date of the test leaves more room for interpreting wage gaps as arising from labor market discrimination.

This finding does not necessarily overturn the conclusions of the Neal-Johnson analysis. At issue is the source of the gap in schooling attainment at the date of the test. The Neal-Johnson premarket factors are a composite of ability and schooling and are likely to reflect both the life cycle experiences and the expectations of the child. To the extent that they reflect expectations of discrimination as embodied in schooling that affects test scores, the scores are contaminated by market discrimination and are not truly premarket factors. An open question is how much of the gap in schooling is due to expectations about future discrimination.

The evidence from data on parents' and children's expectations tells a mixed story. Minority child and parent expectations about the children's schooling prospects are as optimistic at ages 16–17 as those of their white counterparts, although actual schooling outcomes of whites and minorities are dramatically different. Differential expectations at these ages cannot explain the gaps in ability evident in Figure 1.

For children 14 and younger, parent and child expectations about schooling

[17] See Claude Steele & Joshua Aronson, Stereotype Threat and the Test Performance of Academically Successful African Americans, in The Black-White Test Score Gap 401 (Christopher Jencks & Meredith Phillips eds. 1998).

[18] See Paul Sackett, Chaitra Hardison, & Michael Cullen, On Interpreting Stereotype Threat as Accounting for African American–White Differences in Cognitive Tests, 59 Am. Psychologist 7 (2004).

[19] Neal & Johnson, *supra* note 6.

are much lower for blacks than for whites, although only slightly lower for Hispanics than for whites. All groups are still rather optimistic in light of subsequent schooling attendance and performance. At these ages, differences in expectations across groups may lead to differential investments in skill formation. While lower expectations may be a consequence of perceived labor market discrimination, they may also reflect child and parental perception of the lower endowments possessed by minorities, so this evidence is not decisive.

A focus on cognitive skill gaps, while traditional,[20] misses important noncognitive components of social and economic success. We show that noncognitive (behavioral) gaps also open up early. Previous work shows that they play an important role in accounting for market wages. Policies that focus solely on improving cognitive skills miss an important and promising determinant of socioeconomic success and disparity that can be affected by policy.[21]

The rest of the paper proceeds in the following way. Section II presents evidence on the evolution of test score gaps over the life cycle of the child. Section III discusses the evidence on stereotype threat. Section IV presents our evidence on how adjusting for schooling at the date of the test affects the conclusions of the Neal-Johnson analysis and how schooling affects test scores differentially for minorities. Section V discusses our evidence on child and parental expectations. Section VI presents evidence on noncognitive skills that parallels the analysis of Section II. Section VII concludes.

II. Minority-White Differences in Early Test Scores and Early Environments

This section summarizes evidence from the literature and presents original empirical work that demonstrates that minority-white cognitive skill gaps emerge early and persist through childhood and the adolescent years. Christopher Jencks and Meredith Phillips[22] and Greg Duncan and Jeanne Brooks-Gunn[23] document that the black-white test score gap is large for 3- and 4-year-old children. Using the Children of the NLSY79 (CNLSY) survey, a sample of children of the mothers in the 1979 National Longitudinal Survey of Youth data, a variety of studies show that even after controlling for many variables such as individual, family, and neighborhood characteristics, the

[20] See, for example, Christopher Jencks & Meredith Phillips, The Black-White Test Score Gap (1998).

[21] See Pedro Carneiro & James J. Heckman, Human Capital Policy, in Inequality in America: What Role for Human Capital Policies? 77 (James Heckman & Alan Krueger eds. 2003).

[22] Jencks & Phillips, *supra* note 20.

[23] Greg Duncan & Jeanne Brooks-Gunn, Consequences of Growing up Poor (1997).

black-white test score gap is still sizable.[24,25] These studies also document that there are large black-white differences in family environments. Ronald Ferguson[26] summarizes this literature and presents evidence that black children come from much poorer and less educated families than white children, and they are also more likely to grow up in single-parent households. Studies summarized by Ferguson[27] find that the achievement gap is high even for blacks and whites attending high-quality suburban schools.[28] The common finding across these studies is that the black-white gap in test scores is large and that it persists even after one controls for family background variables. Children of different racial and ethnic groups grow up in strikingly different environments.[29] Even after accounting for these environmental factors in a correlational sense, substantial test score gaps remain. Furthermore, these gaps tend to widen with age and schooling: black children show slower measured ability growth with schooling or age than do white children.

This paper presents additional evidence from the children of the persons interviewed in the CNLSY. We have also examined the Early Childhood Longitudinal Survey (ECLS) analyzed by Ferguson[30] and Roland Fryer and Steven Levitt[31] as well as the Children of the Panel Study of Income Dynamics (CPSID) and find similar patterns. We broaden previous analyses to include Hispanic-white differentials. Figure 2 shows the average percentile Peabody Individual Achievement Test (PIAT) Math[32] scores for males in different age groups by race. (Results for females show the same patterns and are available in our Web appendix.[33] For brevity, in this paper we focus only on the male

[24] In a similar study based on the Early Childhood Longitudinal Survey (ECLS), Roland Fryer & Steven Levitt, Understanding the Black-White Test Score Gap in the First Two Years of School, 86 Rev. Econ. Stat. 447 (2004), eliminates the black-white test score gap in math and reading for children at the time they are entering kindergarten, although not in subsequent years. However, the raw test score gaps at ages 3–4 are much smaller in ECLS than in CNLSY and other data sets that have been used to study this issue, so their results are anomalous in the context of a larger literature.

[25] For a description of CNLSY and NLSY79, see Bureau of Labor Statistics, NLS Handbook, 2001 (2001).

[26] Ronald Ferguson, Why America's Black-White School Achievement Gap Persists (unpublished manuscript, Harvard Univ. 2002).

[27] Ronald Ferguson, What Doesn't Meet the Eye: Understanding and Addressing Racial Disparities in High Achieving Suburban Schools (Special Ed., Policy Issues Rep. 2002).

[28] This is commonly referred to as the "Shaker Heights study," although it analyzed many other similar neighborhoods.

[29] See also the discussion in David J. Armor, Maximizing Intelligence (2003).

[30] Ferguson, *supra* note 26.

[31] Fryer & Levitt, *supra* note 24.

[32] Peabody Individual Achievement Test in Mathematics (PIAT Math) measures the child's attainment in mathematics as taught in mainstream education. It consists of 84 multiple-choice questions of increasing difficulty, beginning with recognizing numerals and progressing to geometry and trigonometry. The percentile score was calculated separately for each sex at each age.

[33] Note 9 *supra*.

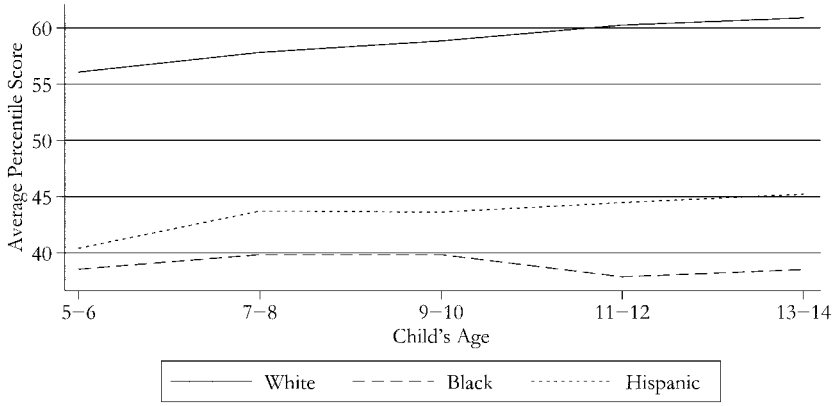

FIGURE 2.—Percentile PIAT Math score by race and age group for CNLSY79 males

results.) Racial and ethnic test score gaps are found as early as ages 5–6 (the earliest ages at which we can measure math scores in CNLSY data).[34] On average, black 5- and 6-year-old boys score almost 18 percentile points below white 5- and 6-year-old boys (that is, if the average white is at the 50th percentile of the test score distribution, the average black is at the 32nd percentile of this distribution). The gap is a bit smaller—16 percent—but still substantial for Hispanics. These findings are duplicated for many other test scores and in other data sets and are not altered if we use median test scores instead of means. Furthermore, as shown in Figure 3, even when we use a test taken at younger ages, racial gaps in test scores can be found at ages 1–2.[35] In general, test score gaps emerge early and persist through adulthood.

For brevity, we focus on means and medians in this paper. However, Figures 1 and 4 illustrate that there is considerable overlap in the distribution of test scores across groups in recent generations. Many black and Hispanic children at ages 5–6 score higher on a math test than the average white child. Statements that we make about medians or means do not apply to all persons in these distributions.

Figure 2 shows that the black-white percentile PIAT Math score gap widens with age. By ages 13–14, the average black is ranked more than 22 percentiles

[34] Instead of using raw scores or standardized scores, we choose to use ranks, or percentiles, since test score scales have no intrinsic meaning. Our results are not sensitive to this procedure.

[35] This is not always the case for women, as shown in our Web appendix (*supra* note 9). The Parts of the Body Test attempts to measure the young child's receptive vocabulary knowledge of orally presented words as a means of estimating intellectual development. The interviewer names each of 10 body parts and asks the child to point to that part of the body. The score is computed by summing the number of correct responses. The percentile score was calculated separately for each sex at each age.

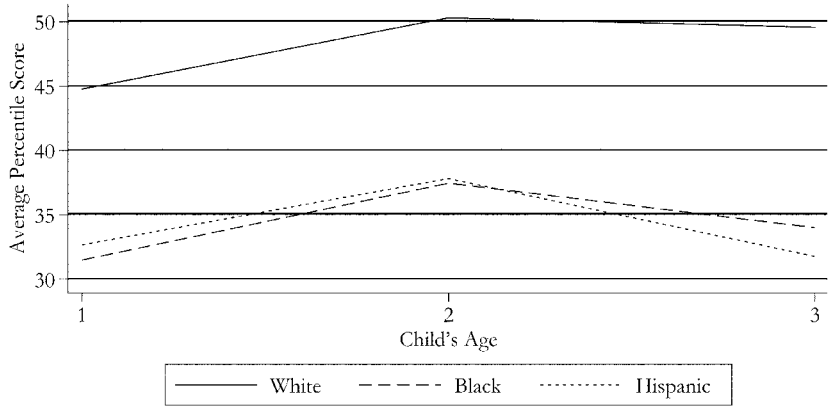

FIGURE 3.—Average percentile Parts of the Body Test score by race and age for CNLSY79 males.

below the average white. In fact, these gaps persist through adulthood. At 13–14, Hispanic boys are almost 16 points below the average white.

When blacks and Hispanics enter the labor market, on average they have a much poorer set of cognitive skills than do whites. Thus, it is not surprising that their average labor market outcomes are so much worse. Furthermore, these skill gaps emerge very early in the life cycle, persist, and, if anything, widen for some groups. Initial conditions (early test scores) are very important since skill begets skill.[36]

The research surveyed by Pedro Carneiro and James Heckman[37] suggests that enhanced cognitive stimulation at early ages is likely to produce lasting gains in achievement test scores in children from disadvantaged environments. If the interventions are early enough, they also appear to raise IQ scores, at least for girls.[38] Home and family environments at early ages, and even the mother's behavior during pregnancy, play crucial roles in the child's development, and black children grow up in environments that are significantly more disadvantaged than those of white children. Figure 5 shows the distributions of long-term or "permanent" family income for blacks, whites, and Hispanics.[39] Minority children are much more likely to grow up in low-income families than are white children. In our Web appendix,[40] we show

[36] See James J. Heckman, Policies to Foster Human Capital, 54 Res. Econ. 3 (2000).

[37] Carneiro & Heckman, *supra* note 21.

[38] See Frances Campbell *et al.*, Early Childhood Education: Young Adult Outcomes from the Abecedarian Project, 6 Applied Developmental Sci. 42 (2002).

[39] Values of permanent income are constructed by taking the average of all nonmissing values of annual family income at ages 0–18 discounted to child's age 0 using a 10 percent discount rate.

[40] Note 9 *supra*.

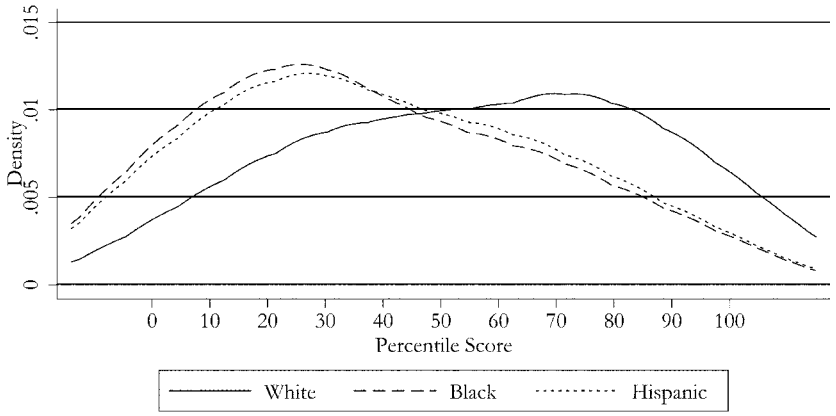

FIGURE 4.—Density of percentile PIAT Math scores at ages 5–6 for CNLSY79 males

that there are also large differences in the level of education and cognitive ability (as measured by the AFQT) of mothers in different ethnic and racial groups (see also Figure 1). Maternal AFQT score is a major predictor of children's test scores.[41] Figure 6 documents that white mothers are much more likely to read to their children at young ages than are minority mothers, and we obtain similar results at other ages.[42] Using this reading variable and other variables in CNLSY such as the number of books, magazines, toys, and musical recordings, family activities (eating together, outings), methods of discipline and parenting, learning at home, television-watching habits, parental expectations for the child (chores, time use), and home cleanliness and safety, we can construct an index of cognitive and emotional stimulation—the home score. This index is always higher for whites than for minorities.[43] The Web appendix also shows that blacks are more likely than whites to grow up in single-parent homes. Hispanics are less likely than blacks to grow up in a single-parent home, although they are much more likely to do so than are whites.

Even after controlling for numerous environmental and family background factors, racial and ethnic test score gaps remain at ages 3–4 for most tests and for virtually all the tests at later ages. Figure 7 shows that, even after

[41] For example, the correlation between percentile PIAT math score and age-corrected maternal AFQT is .4.

[42] See the results for all ages in our Web appendix, *supra* note 9.

[43] As shown in our Web appendix (*id.*), where we document that both cognitive and emotional stimulation indexes are always higher for whites than for blacks at all ages.

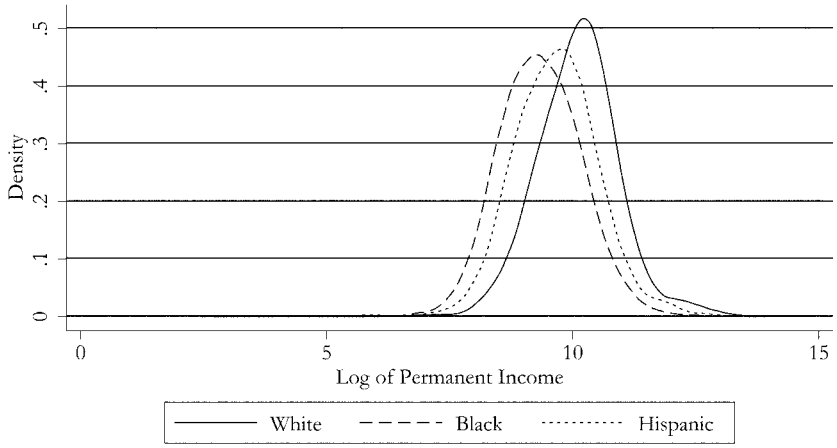

FIGURE 5.—Density of log permanent income for CNLSY79 males and females

adjusting for measures of family background,[44] the black-white gap in percentile PIAT Math scores at ages 5–6 is almost 8 percentile points and at ages 13–14 is close to 11 percentile points. Hispanic-white differentials are reduced more by such adjustments, falling to 7 points at ages 5–6 and to 4 points at ages 13–14. For some tests, differentials frequently are positive or statistically insignificant.[45] Measured home and family environments play an important role in the formation of these skills, although they are not the whole story.[46]

Early test scores for blacks and Hispanics are similar, although Hispanics often perform slightly better. Figure 2 shows that for the PIAT Math score, the Hispanic-black gap is about 2 percentile points.[47] This is much smaller

[44] Scores are adjusted by permanent family income, mother's education, and age-corrected AFQT and home scores. "Adjusted" indicates that we equalized the family background characteristics across all race groups by setting them at the mean to purge the effect of disparities in family environments.

[45] In our Web appendix (*supra* note 9), tables 1A and 1B report that even after controlling for different measures of home environment and child stimulation, the black-white test score gap persists, even though it drops considerably. Results for other tests and other samples can be found in our Web appendix. Even though for some test scores early black-white test score gaps can be eliminated once we control for a large number of characteristics, it is harder to eliminate them at later ages. In the analysis presented here, the most important variable in reducing the test score gap is the mother's cognitive ability, as measured by the AFQT.

[46] However, the home score includes variables such as the number of books, which are clearly choice variables and likely to cause problems in this regression. The variables with the largest effect on the minority-white test score gap are maternal AFQT and raw home score.

[47] The test score is measured in percentile rank The black-white gap is slightly below 18, while the Hispanic-white gap is slightly below 16. This means that the black-Hispanic gap should be around 2.

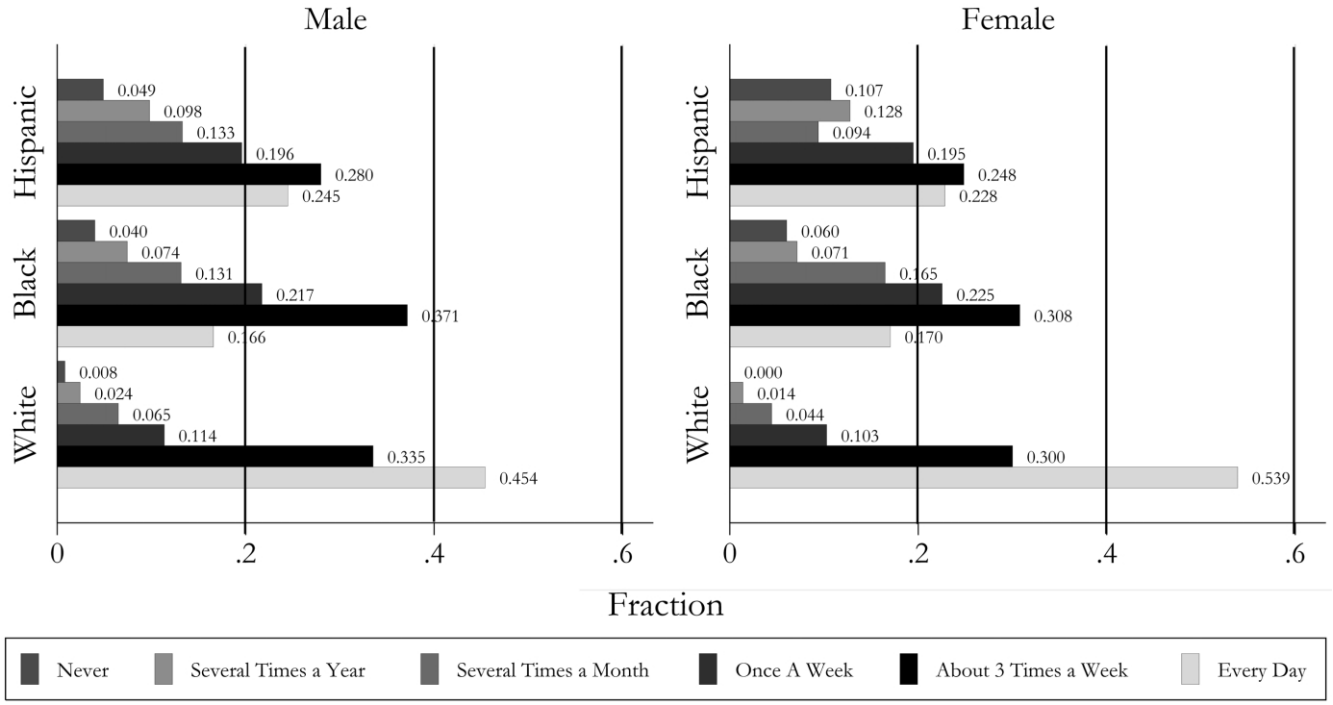

FIGURE 6.—How often mother reads to child at age 2 by race and sex for CNLSY79 males and females. Each bar represents the number of people who report falling in a particular reading frequency cell divided by the total number of people in their race and sex group.

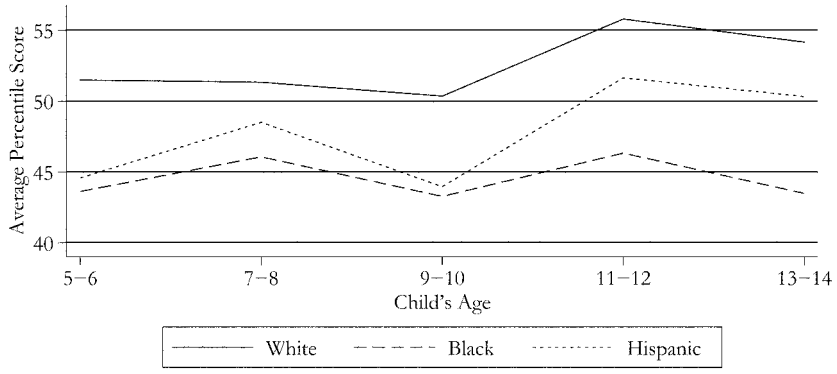

FIGURE 7.—Adjusted percentile PIAT Math score by race and age group for CNLSY79 males.

than either the black-white or the Hispanic-white gap. For the PIAT Math, the black-white gap widens dramatically, especially at later ages, but the Hispanic-white gap does not change substantially with age. For other tests, even when there is some widening of the Hispanic-white gap with age, it tends to be smaller than the widening in the black-white gap in test scores. In particular, when we look at the AFQT scores displayed in Figure 1, which are measured using individuals at ages 16–23, Hispanics clearly have higher scores than do blacks. In contrast, Figure 4 shows a strong similarity between the math scores of blacks and Hispanics at ages 5–6, although there are other tests at which, even at these early ages, Hispanics perform substantially better than blacks. When we control for the effects of home and family environments on test scores, the Hispanic-white test score gap either decreases or is constant over time, while the black-white test score gap tends to widen with age.

III. THE STEREOTYPE THREAT

The fact that substantial racial and ethnic test score gaps open up early in the life cycle of children casts doubt on the empirical importance of the "stereotype threat." It is now fashionable in some circles to attribute gaps in black test scores to racial consciousness on the part of black test takers stemming from the way test scores are used in public discourse to describe minorities.[48] The claim is that blacks perform below their true abilities on standardized tests when a stereotype threat is present. The empirical importance of the stereotype threat in accounting for test score differentials has been greatly overstated in the popular literature.[49] No serious empirical

[48] Steele & Aronson, *supra* note 17.

[49] See the analysis in Sackett, Hardison, & Cullen, *supra* note 18.

scholar assigns any quantitative importance to stereotype threat effects as a major determinant of test score gaps.

Stereotype threats could not have been important when blacks took the first IQ tests at the beginning of the twentieth century, which documented the racial differentials that gave rise to the stereotype. Yet racial IQ gaps are comparable across time.[50] Young children, like the ones studied in this paper, are unlikely to have the heightened racial consciousness about tests and their social significance of the sort claimed to be found by Claude Steele and Joshua Aronson[51] in college students at a few elite universities. Moreover, sizable gaps are found for young Hispanic males—a group for which the stereotype threat remains to be investigated.

Additional evidence on the unimportance of stereotype threat is presented in Table 2.[52] According to the stereotype threat literature, minority test scores understate true ability. If stereotyping affects the test score gap differently across ability levels, the effect of a unit of ability on wages for a black should be different than it is for a white. If the understatement is uniform across all ability levels, the coefficient on a dummy variable for race is overstated in a log wage regression (that is, measured discrimination is understated). If the stereotype threat operates when minorities take the AFQT, their scores should have a different incremental effect on wages than majority AFQT scores.[53] We test this hypothesis using the empirical model in Table 2. We

[50] Charles Murray, The Secular Increase in IQ and Longitudinal Changes in the Magnitude of the Black-White Difference: Evidence from the NLSY (paper presented at the Behavior Genetics Association Meeting, Vancouver 1999), reviews the evidence on the evolution of the black-white IQ gap. In the 1920s—a time when such tests were much more unreliable and black educational attainment much lower—the mean black-white difference was .86 standard deviations. The largest black-white difference appears in the 1960s, with a mean black-white difference of 1.28 standard deviations. The difference ranges from a low of .82 standard deviations in the 1930s to 1.12 standard deviations in the 1970s. However, none of the samples prior to 1960 are nationally representative, and the samples were often chosen so as to effectively bias the black mean upward.

[51] Steele & Aronson, *supra* note 17.

[52] See our Web appendix (*supra* note 9) for evidence on females.

[53] Let $Y = \alpha_0 + \alpha_1 T + \varepsilon$, where $E(\varepsilon \mid T) = 0$. The same equation governs black and white outcomes. The term T is the true test score, and T^* is the test score under stereotype threat:

$$T^* = \gamma_0 + \gamma_1 T + U, \quad E(U \mid T) = 0.$$

Suppose $\mathrm{Cov}(\varepsilon, U) = 0$. Our Web appendix, *supra* note 9, shows that under random sampling, the coefficient on the test score for whites is α_1 and for blacks is

$$\frac{\alpha_1}{\gamma_1}\left(\frac{\gamma_1^2\sigma_T^2}{\gamma_1^2\sigma_T^2 + \sigma_U^2}\right).$$

Intercepts are α_0 and

$$\alpha_0 - \left(\frac{\alpha_1\gamma_0}{\gamma_1}\right)\left(\frac{\gamma_1^2\sigma_T^2}{\gamma_1^2\sigma_T^2 + \sigma_U^2}\right) + \alpha_1 E(T)\left(\frac{\sigma_U^2}{\gamma_1^2\sigma_T^2 + \sigma_U^2}\right),$$

where $E(T)$ is the mean of T, σ_T^2 is the variance of T, and σ_U^2 is the variance of U. Thus, the

TABLE 2

POOLED LOG WAGE REGRESSIONS FOR NLSY MALES, 1990–2000

	1990	1991	1992	1993	1994	1996	1998	2000
Age-corrected AFQT	.1791**	.1714**	.1821**	.1814**	.1966**	.2283**	.2808**	.2462**
	(.0161)	(.0163)	(.0163)	(.0163)	(.0179)	(.0184)	(.0191)	(.0186)
Black	−.045	−.0702*	−.0892**	−.0900**	−.0753*	−.1453**	−.0696	−.0714
	(.0314)	(.0318)	(.0319)	(.0312)	(.0344)	(.0356)	(.0368)	(.0368)
Black × Age-corrected AFQT	.0307	.0014	.0361	.0093	.0228	−.0072	−.0384	.0292
	(.0302)	(.0306)	(.0308)	(.0303)	(.0334)	(.0349)	(.0362)	(.0362)
Hispanic	−.0451	.0116	−.0258	−.0357	−.0073	−.0367	−.0103	−.0645
	(.0326)	(.0330)	(.0329)	(.0327)	(.0364)	(.0381)	(.0389)	(.0386)
Hispanic × Age-corrected AFQT	−.037	−.0629	−.0564	−.0566	−.0529	−.0239	−.0524	−.0345
	(.0329)	(.0336)	(.0336)	(.0339)	(.0375)	(.0398)	(.0401)	(.0398)
Constant	2.2897**	2.2902**	2.3159**	2.3371**	2.3681**	2.4289**	2.4591**	2.5293**
	(.0175)	(.0177)	(.0178)	(.0178)	(.0194)	(.0200)	(.0208)	(.0203)
Observations	1,505	1,514	1,503	1,504	1,485	1,519	1,462	1,404
R^2	.17	.14	.18	.17	.17	.21	.23	.23
F statistic	1.556	1.897	2.787	1.682	1.578	.182	1.139	.927
Significance level	.211	.150	.062	.186	.207	.834	.320	.396

NOTE.—This table reports coefficients of regressions of log hourly wages in each year on the variables listed in the left-hand column. Standard errors are in parentheses. Age-corrected Armed Forces Qualification Test (AFQT) is the standardized residual from the regression of the AFQT score on age at the time of the test dummy variables. AFQT is a subset of four out of 10 Armed Services Vocational Aptitude Battery (ASVAB) tests used by the military for enlistment screening and job assignment. It is the summed score from the word knowledge, paragraph comprehension, mathematics knowledge, and arithmetic reasoning ASVAB tests. All wages are in 1993 dollars. The coefficients on the AFQT variables represent the effect of a 1-standard-deviation increase in the score on the log hourly wage. Since the wage is measured in log points, the gaps for blacks and Hispanics correspond approximately to percentage point differences relative to the white mean; that is, the black-white gap of −.25 in 1990 corresponds to wages for blacks in that year that are 25% lower than those for whites. The joint null hypothesis for the Wald F test was that all of the race × AFQT interaction coefficients are zero. NLSY = National Longitudinal Survey of Youth.

* Significant at the 5% level.

** Significant at the 1% level.

estimate the effect of black and Hispanic AFQT scores relative to the effect of white AFQT scores on log wages as extracted from the NLSY79. This amounts to testing for racial AFQT interactions in a log wage equation. While there is some (weak) evidence that black scores have a larger effect on log wages than white scores, the black-AFQT interaction coefficients are small in magnitude and imprecisely determined. For Hispanics, the estimated AFQT interaction coefficients are negative and, again, not precisely determined. In our Web appendix, we also graph the mean log wage by AFQT decile by race. There is no particular pattern of convergence or divergence across ability levels when evaluated over common supports.

The stereotype literature substitutes wishful thinking for substantial evidence. There is no evidence that it accounts for an important fraction of minority-white test score gaps or that test scores are not good measures of productivity.[54]

IV. The Differential Effect of Schooling on Test Scores

We have established that cognitive test scores are correlated with home and family environments and that test score gaps increase with age and schooling. The research of Karsten Hansen, James Heckman, and Kathleen Mullen[55] and Heckman, Larenas, and Urzua[56] shows that the AFQT scores used by Neal and Johnson[57] are affected by the schooling attainment of individuals at the time they take the test. Therefore, one reason for the divergence of black and white test scores over time may be differential schooling attainments. Figure 8 shows the schooling completed at the test date for the six demographic groups in the age ranges of the NLSY used by Neal and Johnson. Blacks have completed (slightly) less schooling at the test date than whites but substantially more than Hispanics.

Table 3 presents estimates of the effect of schooling at test date on AFQT scores for individuals in different demographic groups in the NLSY, using a version of the nonparametric method developed by Hansen, Heckman, and Mullen.[58] Their method isolates the causal effect of schooling attained at the

---

intercepts for blacks are upward biased. The slope for blacks in general may be greater than or less than α_1, depending on whether the gap widens with T ($\gamma_1 < 1$) or shrinks ($\gamma_1 > 1$). When $\sigma_U^2 = 0$ ($U = 0$) and $\gamma_1 = 1$, the slopes are the same for blacks and whites, but the intercepts for blacks are upward biased. This method underestimates the amount of discrimination.

[54] A circular version of the stereotype threat argument would claim that minorities also underperform at the workplace because of stereotype threat there, so using measured wages to capture productivity understates true black productivity. This form of the stereotype threat argument is irrefutable. All measures are contaminated.

[55] Hansen, Heckman, & Mullen, *supra* note 14.

[56] Heckman, Larenas, & Urzua, *supra* note 16.

[57] Neal & Johnson, *supra* note 6.

[58] Hansen, Heckman, & Mullen, *supra* note 14. Heckman, Larenas, & Urzua, *supra* note 16, presents a more refined analysis of the racial/ethnic wage gap using the analysis of Hansen, Heckman, & Mullen that supports all of our main conclusions. See also the note to Table 3.

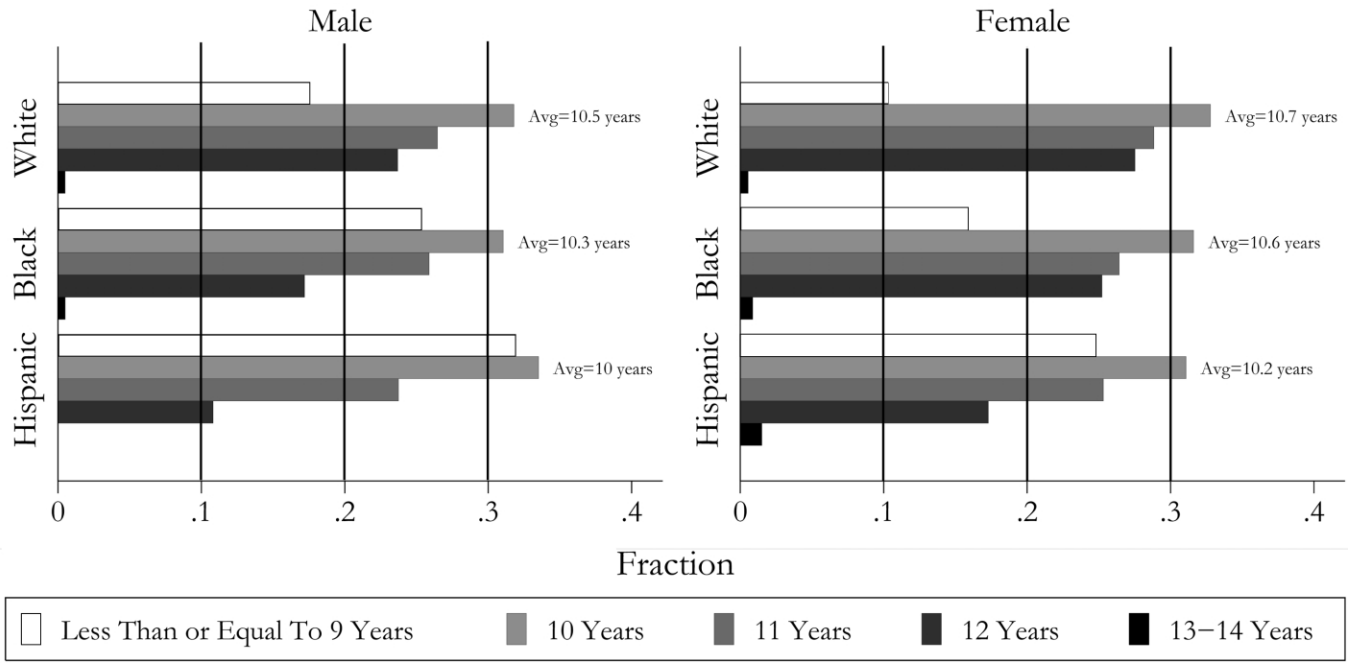

FIGURE 8.—Highest category of schooling completed at the test date by race, sex, and age for NLSY79 males and females born after 1961. Each bar represents the number of people who report falling in a particular reading frequency cell divided by the total number of people in their race and sex group.

TABLE 3

EFFECT OF YEARS OF SCHOOLING ON AFQT SCORES FOR INDIVIDUALS IN NLSY79

	WHITE		BLACK		HISPANIC	
SCHOOLING AT TEST DATE	Males	Females	Males	Females	Males	Females
9 years:	12.68	14.66	5.21	5.95	9.80	11.22
S.E.	(1.51)	(1.70)	(1.58)	(1.74)	(1.90)	(1.79)
N	378	343	234	234	167	179
10 years:	16.94	16.12	9.22	5.65	16.33	14.99
S.E.	(1.52)	(1.68)	(1.53)	(1.73)	(1.94)	(1.85)
N	377	368	282	237	161	169
11 years:	22.02	18.55	8.87	10.62	18.67	16.83
S.E.	(1.54)	(1.72)	(1.58)	(1.75)	(2.11)	(2.01)
N	366	327	247	243	120	122
12 years:	23.12	21.05	11.96	11.12	21.29	16.46
S.E.	(1.49)	(1.64)	(1.58)	(1.70)	(2.08)	(1.95)
N	630	758	322	393	171	198
13–15 years:	26.60	24.37	15.37	14.23	23.96	18.92
S.E.	(1.73)	(1.83)	(2.23)	(2.00)	(2.82)	(2.55)
N	266	326	98	180	72	81
16 or more years	29.02	25.71	28.43	22.66	33.11	31.18
S.E.	(2.13)	(2.24)	(3.56)	(2.99)	(4.70)	(4.03)
N	108	103	27	34	17	22
Mean schooling-corrected AFQT, 8 years of schooling:						
Mean	52.50	51.95	36.79	37.44	38.45	36.87
S.D.	(19.11)	(17.81)	(17.76)	(16.15)	(19.17)	(17.80)
N	165	110	172	98	154	137

NOTE.—Coefficients of a regression of raw Armed Forces Qualification Test (AFQT) scores on schooling at test date and completed schooling are reported. Completed schooling is included to control for unobserved ability that may be correlated with schooling at the test date (control function). Schooling at the test date is measured by the set of dummy variables reported in the top section of the table. The omitted category is 8 years of schooling or less. Completed schooling (not reported) is measured with dummy variables for four categories: less than high school, high school, some college, and college. We run a different regression for each race and gender group. The numbers reported in the bottom section are the predicted value of these regressions when schooling at the test date is equal to 8 years of schooling or less. For example, white males with 12 years of schooling at the test date would score 23.12 points higher on the AFQT than they would if they had only 8 years of schooling at that time. *N* = number of observations. S.E. = standard error; S.D. = standard deviation.

test date on test scores controlling for unobserved factors that lead to selective differences in schooling attainment. This table shows that the effect of schooling on test scores is much larger for whites and Hispanics than it is for blacks over most ranges of schooling. As a result, even though Hispanics have fewer years of completed schooling than blacks at the time they take the AFQT, on average Hispanics score better on the AFQT than do blacks.

There are different explanations for these findings. Carneiro and Heckman, Cunha and Heckman, and Cunha and colleagues[59] suggest that one important

[59] Carneiro & Heckman, *supra* note 21; Flavio Cunha & James Heckman, The Technology of Skill Formation (unpublished manuscript, Univ. Chicago 2004); Flavio Cunha *et al.*, Interpreting the Evidence on Life Cycle Skill Formation, in Handbook of Education Economics (Finis Welch & Eric Hanushek eds., forthcoming 2005).

feature of the learning process is complementarity and self-productivity between initial endowments of human capital and subsequent learning.[60] Higher levels of human capital raise the productivity of learning.[61] Since minorities and whites start school with very different initial conditions, their learning paths can diverge dramatically over time. A related explanation may be that blacks and nonblacks learn at different rates because blacks attend lower quality schools than whites.[62]

Janet Currie and Duncan Thomas[63] show that test score gains of participants in the Head Start program tend to fade completely for blacks but not for whites. They suggest that one reason may be that blacks attend worse schools than whites, and therefore blacks are not able to maintain initial test score gains. Both early advantages and disadvantages as well as school quality are likely to be important factors in the human capital accumulation process.

In light of the greater growth in test scores of Hispanics that is parallel to that of whites, explanations based on schooling quality are not entirely compelling. Hispanics start from similar initial disadvantages in family environments and face school and neighborhood environments similar to those faced by blacks.[64] They also have early levels of test scores similar to those found in the black population.[65]

To analyze the consequences of correcting for different levels of schooling at the test date, we reanalyze the Neal-Johnson[66] data using AFQT scores corrected for the race- or ethnicity-specific effect of schooling while equalizing the years of schooling attained at the date of the test across all racial/ethnic groups. The results of this adjustment are presented in Table 4. This adjustment is equivalent to replacing each individual's AFQT score by the score we would measure if he or she would have stopped his or her formal education after eighth grade.[67] In other words, we use eighth-grade-adjusted AFQT scores for everyone. Since the effect of schooling on test scores is

[60] For example, see the model in Yoram Ben-Porath, The Production of Human Capital and the Life Cycle of Earnings, 75 J. Pol. Econ. 352 (1967). See also Cunha *et al.*, *supra* note 59.

[61] See the evidence in James Heckman, Lance Lochner, & Christopher Taber, Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents, 1 Rev. Econ. Dynamics 1 (1998).

[62] Cunha & Heckman, *supra* note 59, shows that complementarity implies that early human capital increase the productivity of later investments in human capital and that early investments that are not followed up by later investments in human capital are not productive.

[63] Janet Currie & Duncan Thomas, School Quality and the Longer-Term Effects of Head Start, 35 J. Hum. Resources 755 (2000).

[64] The evidence for CNLSY is presented in our Web appendix (*supra* note 9).

[65] Heckman, Larenas, & Urzua, *supra* note 16, presents a more formal analysis of the effect of schooling quality on test scores, showing that schooling inputs explain little of the differential growth in test scores among blacks, whites, and Hispanics.

[66] Neal & Johnson, *supra* note 6.

[67] However, the score is affected by attendance in kindergarten, 8 further years of schooling, and any school quality differentials in those years.

TABLE 4

CHANGE IN THE BLACK-WHITE LOG WAGE GAP INDUCED BY CONTROLLING FOR SCHOOLING-CORRECTED AFQT SCORES, 1990–2000

	1990		1991		1992		1993		1994		1996		1998		2000	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Males:																
Black	−.250	−.133	−.251	−.149	−.302	−.180	−.282	−.171	−.286	−.165	−.373	−.230	−.333	−.160	−.325	−.172
	(.028)	(.029)	(.028)	(.029)	(.029)	(.029)	(.028)	(.029)	(.031)	(.031)	(.032)	(.032)	(.034)	(.034)	(.035)	(.033)
Hispanic	−.174	−.070	−.113	−.013	−.146	−.044	−.159	−.058	−.143	−.029	−.186	−.067	−.195	−.047	−.215	−.088
	(.032)	(.032)	(.032)	(.032)	(.033)	(.032)	(.032)	(.032)	(.036)	(.036)	(.038)	(.038)	(.040)	(.038)	(.040)	(.038)
Age	. . .	.065	. . .	.043	. . .	.055	. . .	.045	. . .	.040	. . .	.039	. . .	.036	. . .	.029
	. . .	(.014)	. . .	(.015)	. . .	(.015)	. . .	(.015)	. . .	(.016)	. . .	(.017)	. . .	(.017)	. . .	(.017)
AFQT	. . .	.153	. . .	.131	. . .	.155	. . .	.144	. . .	.159	. . .	.184	. . .	.221	. . .	.211
	. . .	(.013)	. . .	(.013)	. . .	(.013)	. . .	(.013)	. . .	(.014)	. . .	(.015)	. . .	(.015)	. . .	(.015)
$AFQT^2$	. . .	.001	. . .	.009	. . .	.015	. . .	.017	. . .	.028	. . .	.031	. . .	.040	. . .	.036
	. . .	(.010)	. . .	(.010)	. . .	(.010)	. . .	(.010)	. . .	(.011)	. . .	(.012)	. . .	(.012)	. . .	(.012)
Constant	2.375	.540	2.372	1.085	2.404	.732	2.423	.982	2.458	1.119	2.533	1.140	2.589	1.172	2.629	1.425
	(.017)	(.392)	(.017)	(.412)	(.017)	(.426)	(.017)	(.437)	(.018)	(.493)	(.019)	(.547)	(.020)	(.603)	(.020)	(.628)
N	1,538	1,505	1,553	1,514	1,536	1,503	1,542	1,504	1,522	1,485	1,554	1,519	1,494	1,462	1,438	1,404
Females:																
Black	−.172	−.045	−.200	−.066	−.201	−.083	−.167	−.020	−.148	−.014	−.147	.025	−.201	−.043	−.200	−.041
	(.031)	(.031)	(.032)	(.033)	(.031)	(.031)	(.035)	(.035)	(.035)	(.035)	(.035)	(.034)	(.034)	(.034)	(.036)	(.035)
Hispanic	−.003	.116	−.017	.107	−.059	.069	.009	.145	−.018	.119	−.006	.149	−.069	.098	−.064	.096
	(.035)	(.035)	(.037)	(.037)	(.036)	(.035)	(.039)	(.039)	(.040)	(.040)	(.041)	(.039)	(.039)	(.038)	(.041)	(.040)
Age	. . .	.015	. . .	.042	. . .	.021	. . .	.024	. . .	.015	. . .	−.003	. . .	.019	. . .	−.010
	. . .	(.016)	. . .	(.017)	. . .	(.016)	. . .	(.018)	. . .	(.018)	. . .	(.018)	. . .	(.017)	. . .	(.018)
AFQT	. . .	.188	. . .	.197	. . .	.187	. . .	.221	. . .	.221	. . .	.245	. . .	.228	. . .	.235
	. . .	(.016)	. . .	(.017)	. . .	(.016)	. . .	(.018)	. . .	(.018)	. . .	(.018)	. . .	(.017)	. . .	(.018)
$AFQT^2$	. . .	.010	. . .	.009	. . .	.010	. . .	.006	. . .	−.008	. . .	.022	. . .	.017	. . .	.005
	. . .	(.013)	. . .	(.014)	. . .	(.013)	. . .	(.015)	. . .	(.015)	. . .	(.015)	. . .	(.015)	. . .	(.015)
Constant	2.141	1.633	2.175	.893	2.193	1.488	2.174	1.337	2.218	1.662	2.246	2.228	2.311	1.550	2.339	2.608
	(.019)	(.424)	(.020)	(.472)	(.019)	(.465)	(.021)	(.530)	(.022)	(.565)	(.022)	(.588)	(.021)	(.612)	(.022)	(.671)
N	1,356	1,325	1,335	1,299	1,317	1,278	1,319	1,281	1,318	1,287	1,381	1,343	1,370	1,328	1,316	1,276

NOTE.—This table reports coefficients of regressions of log hourly wages in each year on the variables listed in the left-hand column. Values given are for participants in the National Longitudinal Survey of Youth born after 1961. Schooling-corrected Armed Forces Qualification Test (AFQT) is the standardized residual from the regression of the AFQT score on age at the time of the test dummy variables and final level of schooling completed during lifetime. AFQT is a subset of four out of 10 Armed Services Vocational Aptitude Battery (ASVAB) tests used by the military for enlistment screening and job assignment. It is the summed score from the word knowledge, paragraph comprehension, mathematics knowledge, and arithmetic reasoning ASVAB tests. All wages are in 1993 dollars.

higher for whites than for blacks, and whites have more schooling than blacks at the date of the test, this adjustment reduces the test scores of whites much more than those for blacks. The black-white male wage gap is cut only in half (as opposed to 76 percent) when we use this new measure of skill, and a substantial unexplained residual remains. The adjustment has little effect on the Hispanic-white wage gap, but a wage gap for black women emerges when using the schooling-adjusted measure that did not appear in the original Neal-Johnson study.

Adjusting for schooling at the date of the test reduces the test score gap. This evidence raises the larger question of what a premarket factor is. Neal and Johnson do not condition on schooling in explaining black-white wage gaps, arguing that schooling is affected by expectations of adverse market opportunities facing minorities and that conditioning on such a contaminated variable would spuriously reduce the estimated wage gap. We present direct evidence on this claim below.

Their reasoning is not entirely coherent. If expectations of discrimination affect schooling, the very logic of their premarket argument suggests that they should control for the impact of schooling on test scores before using test scores to measure premarket factors. Neal and Johnson[68] assume that schooling at the time the test is taken is not affected by expectations of discrimination in the market, while later schooling is.

This distinction is arbitrary. A deeper investigation of the expectation formation process and feedback is required. One practical conclusion with important implications for interpretation of the evidence is that the magnitude of the wage gap one can eliminate by performing a Neal-Johnson analysis depends on the age at the time the test is taken. We find that the earlier the test is taken, the smaller the unadjusted test score gap, and the larger the fraction of the wage gap that is unexplained by the residual. Figure 9 shows how adjusting measured ability for schooling at the time of the test at different levels of attained schooling affects the adjusted wage gap for black males. For example, the log wage gap that we obtain when using eleventh-grade test scores corresponds to that using an AFQT correction equal to 11. The later the grade at which we adjust the test score, the lower the estimated gap. This is because a test score gap opens up at later schooling levels, and hence adjustment reduces the gap by a larger amount at later schooling levels.[69]

[68] Neal & Johnson, *supra* note 6.

[69] The figure omits the results for the 16-and-over category because the low number of minorities makes correction of test scores to that level much less reliable than correction to the other schooling levels. The unadjusted line shows the black-white log wage gap we observe if we do not depend on the grade to which we are correcting the test score. The adjusted line shows the black-white log wage gap after we adjust for the AFQT scores corrected to different grades. In our Web appendix (*supra* note 9), we present the same analysis for females and Hispanics.

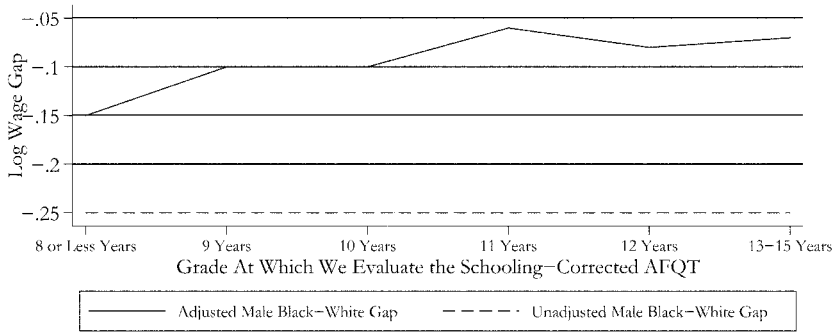

FIGURE 9.—Residual black-white log wage gap in 1991 by grade at which we evaluate schooling-corrected AFQT scores for NLSY79 males.

Finally we show that adjusting for "expectations-contaminated" completed schooling by entering it as a direct regression in a log wage equation does not operate in the fashion conjectured by Neal and Johnson. Table 5 shows that when we adjust wage differences for completed schooling as well as schooling-adjusted AFQT, wage gaps widen relative to the simple adjustment. This runs contrary to the simple intuition that schooling embodies expectations of market discrimination, so conditioning on it will eliminate wage gaps.[70] The deeper issue, not resolved in this paper or the literature, is what productivity factors to condition on in measuring discrimination. Schooling and measured ability are both valid candidate productivity variables. Conditioning on them singly or jointly and eliminating spurious endogeneity effects produces conceptually different measures of the wage gap, all of which answer distinct but economically interesting questions. Both variables may be affected by discrimination. Looking only at outcome equations, one cannot settle what is a productivity characteristic and what is contaminated and what is not.[71,72] Deleting potential contaminated variables does not, in general, produce the conceptually desired measure of discrimination.

[70] The simple intuition, however, can easily be shown to be wrong, so the evidence in these tables is not decisive on the presence of discrimination in the labor market. The basic idea is that if both schooling and the test score are correlated with an unmeasured discrimination component in the error term, the bias for the race dummy may be either positive or negative depending on the strength of the correlation among the contaminated variables and their correlation with the error term. See the discussion in our Web appendix (*id.*), where we show that if both schooling and test score are correlated with factors leading to discrimination in earnings, the estimated discrimination effect may be upward or downward biased by adding schooling as a regressor.

[71] See Robert Bornholz & James J. Heckman, Measuring Disparate Impacts and Extending Disparate Impact Doctrine to Organ Transplantation, 48 Persp. Biology & Med. S95 (2005).

[72] As pointed out to us by an anonymous referee, another reason for excluding years of schooling from the log wage equation is that schooling overstates the amount of human capital black children receive relative to white children, say because of differential schooling quality.

TABLE 5

CHANGE IN THE BLACK-WHITE LOG WAGE GAP INDUCED BY CONTROLLING FOR SCHOOLING-CORRECTED AFQT SCORES AND HIGHEST GRADE COMPLETED, 1990–2000

	1990		1991		1992		1993		1994		1996		1998		2000	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Males:																
Black	−.250	−.144	−.251	−.158	−.302	−.189	−.282	−.182	−.286	−.175	−.373	−.241	−.333	−.175	−.325	−.194
	(.028)	(.028)	(.028)	(.028)	(.029)	(.028)	(.028)	(.028)	(.031)	(.031)	(.032)	(.031)	(.034)	(.032)	(.035)	(.032)
Hispanic	−.174	−.056	−.113	.005	−.146	−.032	−.159	−.044	−.143	−.018	−.186	−.056	−.195	−.040	−.215	−.070
	(.032)	(.032)	(.032)	(.032)	(.033)	(.032)	(.032)	(.032)	(.036)	(.035)	(.038)	(.036)	(.040)	(.037)	(.040)	(.037)
Age	. . .	.062	. . .	.041	. . .	.052	. . .	.042	. . .	.036	. . .	.033	. . .	.031	. . .	.024
	. . .	(.014)	. . .	(.014)	. . .	(.014)	. . .	(.014)	. . .	(.016)	. . .	(.016)	. . .	(.017)	. . .	(.016)
AFQT	. . .	.096	. . .	.079	. . .	.097	. . .	.082	. . .	.098	. . .	.093	. . .	.130	. . .	.119
	. . .	(.015)	. . .	(.015)	. . .	(.015)	. . .	(.015)	. . .	(.016)	. . .	(.017)	. . .	(.017)	. . .	(.017)
$AFQT^2$	. . .	−.015	. . .	−.005	. . .	−.001	. . .	−.001	. . .	.010	. . .	.005	. . .	.013	. . .	.008
	. . .	(.010)	. . .	(.010)	. . .	(.010)	. . .	(.010)	. . .	(.011)	. . .	(.012)	. . .	(.012)	. . .	(.012)
HGC	. . .	.044	. . .	.040	. . .	.043	. . .	.047	. . .	.046	. . .	.065	. . .	.066	. . .	.066
	. . .	(.006)	. . .	(.006)	. . .	(.006)	. . .	(.006)	. . .	(.006)	. . .	(.006)	. . .	(.007)	. . .	(.006)
Constant	2.375	.113	2.372	.668	2.404	.281	2.423	.501	2.458	.679	2.533	.540	2.589	.549	2.629	.802
	(.017)	(.390)	(.017)	(.410)	(.017)	(.422)	(.017)	(.431)	(.018)	(.488)	(.019)	(.531)	(.020)	(.586)	(.020)	(.608)
N	1,538	1,504	1,553	1,513	1,536	1,503	1,542	1,504	1,522	1,485	1,554	1,519	1,494	1,462	1,438	1,404

Females:																
Black	−.172	−.081	−.200	−.101	−.201	−.131	−.167	−.073	−.148	−.069	−.147	−.035	−.201	−.088	−.200	−.086
	(.031)	(.030)	(.032)	(.032)	(.031)	(.030)	(.035)	(.033)	(.035)	(.034)	(.035)	(.033)	(.034)	(.032)	(.036)	(.034)
Hispanic	−.003	.120	−.017	.111	−.059	.073	.009	.139	−.018	.118	−.006	.137	−.069	.091	−.064	.096
	(.035)	(.033)	(.037)	(.036)	(.036)	(.033)	(.039)	(.037)	(.040)	(.038)	(.041)	(.037)	(.039)	(.036)	(.041)	(.038)
Age	. . .	.013	. . .	.036	. . .	.013	. . .	.014	. . .	.006	. . .	−.009	. . .	.013	. . .	−.017
	. . .	(.015)	. . .	(.016)	. . .	(.015)	. . .	(.017)	. . .	(.017)	. . .	(.017)	. . .	(.017)	. . .	(.017)
AFQT	. . .	.106	. . .	.113	. . .	.094	. . .	.121	. . .	.111	. . .	.130	. . .	.119	. . .	.127
	. . .	(.017)	. . .	(.019)	. . .	(.017)	. . .	(.019)	. . .	(.020)	. . .	(.019)	. . .	(.019)	. . .	(.020)
AFQT^2	. . .	−.001	. . .	−.002	. . .	−.007	. . .	−.009	. . .	−.026	. . .	.003	. . .	.004	. . .	−.007
	. . .	(.013)	. . .	(.014)	. . .	(.013)	. . .	(.014)	. . .	(.015)	. . .	(.014)	. . .	(.014)	. . .	(.014)
HGC	. . .	.063	. . .	.064	. . .	.075	. . .	.075	. . .	.081	. . .	.081	. . .	.076	. . .	.073
	. . .	(.006)	. . .	(.006)	. . .	(.006)	. . .	(.007)	. . .	(.007)	. . .	(.006)	. . .	(.006)	. . .	(.006)
Constant	2.141	.913	2.175	.270	2.193	.784	2.174	.720	2.218	.932	2.246	1.397	2.311	.802	2.339	1.928
	(.019)	(.413)	(.020)	(.460)	(.019)	(.443)	(.021)	(.508)	(.022)	(.538)	(.022)	(.558)	(.021)	(.583)	(.022)	(.643)
N	1,356	1,325	1,335	1,299	1,317	1,278	1,319		1,318	1,286	1,381	1,343	1,370	1,328	1,316	1,276

NOTE.—This table reports coefficients of regressions of log hourly wages in each year on the variables listed in the left-hand column. Values given are for participants in the National Longitudinal Survey of Youth born after 1961. Schooling-corrected Armed Forces Qualification Test (AFQT) is the standardized residual from the regression of the AFQT score on age at the time of the test dummy variables and final level of schooling completed during lifetime. AFQT is a subset of four out of 10 Armed Services Vocational Aptitude Battery (ASVAB) tests used by the military for enlistment screening and job assignment. It is the summed score from the word knowledge, paragraph comprehension, mathematics knowledge, and arithmetic reasoning ASVAB tests. All wages are in 1993 dollars. HGC is the highest observed level of schooling completed during the individual's lifetime.

Ours is a worst-case analysis for the Neal-Johnson study.[73] If we assign all racial and ethnic schooling differences to expectations of discrimination in the labor market, the results for blacks are less sharp than Neal and Johnson claim. Yet even in the worst-case scenario, adjusting for ability corrected for schooling and schooling as a direct effect on wages substantially reduces minority-majority wage gaps over the unadjusted case. The evidence presented in Section II about the early emergence of ability differentials is reinforced by the early emergence of differential grade repetition gaps for minorities documented by Cameron and Heckman.[74] Most of the schooling gap at the date of the test emerges in the early years at ages when child expectations about future discrimination are unlikely to be operative. One might argue that these early schooling and ability gaps are due to parental expectations of poor labor markets for minority children. We next examine data on child and parental expectations.

V. The Role of Expectations

The argument that minority children perform worse on tests because they expect to be less well rewarded in the labor market than whites for the same test score or schooling level is implausible because expectations of labor market rewards are unlikely to affect the behavior of children as early as ages 3 or 4, when test score gaps are substantial across different ethnic and racial groups. The argument that minorities invest less in skills because both minority children and minority parents have low expectations about their performance in school and in the labor market has mixed empirical backing.

Data on expectations are hard to find, and when they are available they are often difficult to interpret. For example, in the NLSY97, black 17- and 18-year-olds report that the probability of dying next year is 22 percent, while whites report a probability of dying of 16 percent.[75] Both numbers are absurdly high. Minorities usually report higher expectations than whites of committing a crime, being incarcerated, and being dead next year, and these adverse expectations may reduce their investment in human capital. Expectations reported by parents and children for the adolescent years for a variety of outcomes are given in our Web appendix.[76]

Schooling expectations measured in the late teenage years are very similar for minorities and whites. They are slightly lower for Hispanics. Table 6

If this effect is strong enough, including years of schooling will overstate the racial wage differential. Table 3 shows that years of schooling for black children have less effect on human capital (the test score) than years of schooling for white children. However, Heckman, Larenas, & Urzua, *supra* note 16, shows that measured schooling quality accounts for little of the gap or the growth in the gap between blacks and whites.

[73] Neal & Johnson, *supra* note 6.

[74] Cameron & Heckman, *supra* note 11.

[75] See our Web appendix, table 3, for evidence on expectations from NLSY97 (*supra* note 9).

[76] *Id.*

TABLE 6

JUVENILE EXPECTATIONS ABOUT SCHOOL ENROLLMENT IN 1998: NLSY97 MALES

	BLACK		HISPANIC		WHITE	
	Expected	Actual	Expected	Actual	Expected	Actual
All individuals	.912	.734	.881	.717	.934	.790
	(.232)	(.442)	(.265)	(.451)	(.219)	(.407)
Individuals enrolled in 1997	.936	.764	.915	.758	.957	.819
	(.188)	(.425)	(.217)	(.429)	(.173)	(.385)

NOTE.—In NLSY97, round 1, respondents who were born in 1980 or 1981 were surveyed on their beliefs about the future. Asked to assess the probability that certain events would occur in a specified time period, the respondents were instructed to use a scale from 0 (impossible) to 100 (certain). In the Expected columns, we report the percentage of each race group that expects to be enrolled in the next year. In the Actual columns, we report the percentage of each race group that is actually enrolled in that year. Expectations were measured at ages 17–18.

reports the mean expected probability of being enrolled in school next year for black, white, and Hispanic 17- and 18-year-old males. Among those individuals enrolled in 1997, on average whites expect to be enrolled next year with 95.7 percent probability. Blacks expect that they will be enrolled next year with a 93.6 percent probability. Hispanics expect to be enrolled with a 91.5 percent probability. If expectations about the labor market are adverse for minorities, they should translate into adverse expectations for the child's education. Yet these data do not reveal this. Moreover, all groups substantially overestimate actual enrollment probabilities. The difference in expectations between blacks and whites is very small and is less than half the difference in actual (realized) enrollment probabilities (81.9 percent for whites versus 76.4 percent for blacks). The gap is wider for Hispanics. Table 7 reports parental schooling expectations for white, black, and Hispanic males for the same individuals used to compute the numbers in Table 6. It shows that, conditional on being enrolled in 1997 (the year the expectation question is asked), black parents expect their sons to be enrolled next year with a 90.9 percent probability, while for whites this expectation is 95.4 percent. For Hispanics, this number is lower (88.5 percent) but still substantial. Parents overestimate enrollment probabilities for their sons, but black parents have lower expectations than white parents. For females, the racial and ethnic differences in parental expectations are smaller than those for males.[77]

For expectations measured at earlier ages the story is dramatically different. Figures 10 and 11 show that, for the CNLSY group, both black and Hispanic children and their parents have more pessimistic expectations about schooling than do white children, and more pessimistic expectations may lead to lower

[77] *Id.*

TABLE 7

PARENTAL EXPECTATIONS ABOUT YOUTH SCHOOL ENROLLMENT IN 1998: NLSY79 MALES

	BLACK		HISPANIC		WHITE	
	Expected	Actual	Expected	Actual	Expected	Actual
All individuals	.885	.734	.880	.717	.930	.790
	(.255)	(.442)	(.259)	(.451)	(.217)	(.407)
Individuals enrolled in 1997	.909	.764	.911	.758	.954	.819
	(.221)	(.425)	(.220)	(.429)	(.169)	(.385)

NOTE.—In round 1, parents of NLSY97 respondents who were born in 1980 or 1981 were surveyed on their beliefs about their children's future. Asked to assess the probability that certain events would occur in a specified time period, the respondents were instructed to use a scale from 0 (impossible) to 100 (certain). In the Expected columns, we report the percentage of each race group that expects its children to be enrolled in the next year. In the Actual columns, we report the percentage of each race group that is actually enrolled in that year. Expectations were measured at ages 17–18.

investments in skills, less effort in schooling, and lower levels of ability. These patterns are also found in the CPSID and ECLS groups.[78]

If the more pessimistic expectations of minorities are a result of perceived market discrimination, then lower levels of investment in children that translate into lower levels of ability and skill at later ages are attributable to market discrimination. Ability would not be a premarket factor. However, lower expectations for minorities may not be a result of discrimination but just a rational response to the fact that minorities do not do as well in school as whites. This may be due to environmental factors unrelated to expectations of discrimination in the labor market. Whether this phenomenon itself is a result of discrimination is an open question. Expectation formation models are very complex and often lead to multiple equilibria and therefore are difficult to test empirically. However, the evidence reported here does not provide much support for the claim that the ability measure used by Neal and Johnson[79] is substantially contaminated by expectational effects.

VI. THE EVIDENCE ON NONCOGNITIVE SKILLS

Controlling for scholastic ability in accounting for minority-majority wage gaps captures only part of the endowment differences between groups but receives most of the emphasis in the literature on black-white gaps in wages. An emerging body of evidence, summarized by Samuel Bowles, Herbert Gintis, and Melissa Osborne,[80] Carneiro and Heckman,[81] and Heckman, Stix-

[78] For CNLSY teenagers, expectations across racial groups seem to converge at later ages. See our Web appendix (*id.*).

[79] Neal & Johnson, *supra* note 6.

[80] Samuel Bowles, Herbert Gintis, & Melissa Osborne, The Determinants of Earnings: A Behavioral Approach, 39 J. Econ. Literature 1137 (2001).

[81] Carneiro & Heckman, *supra* note 21.

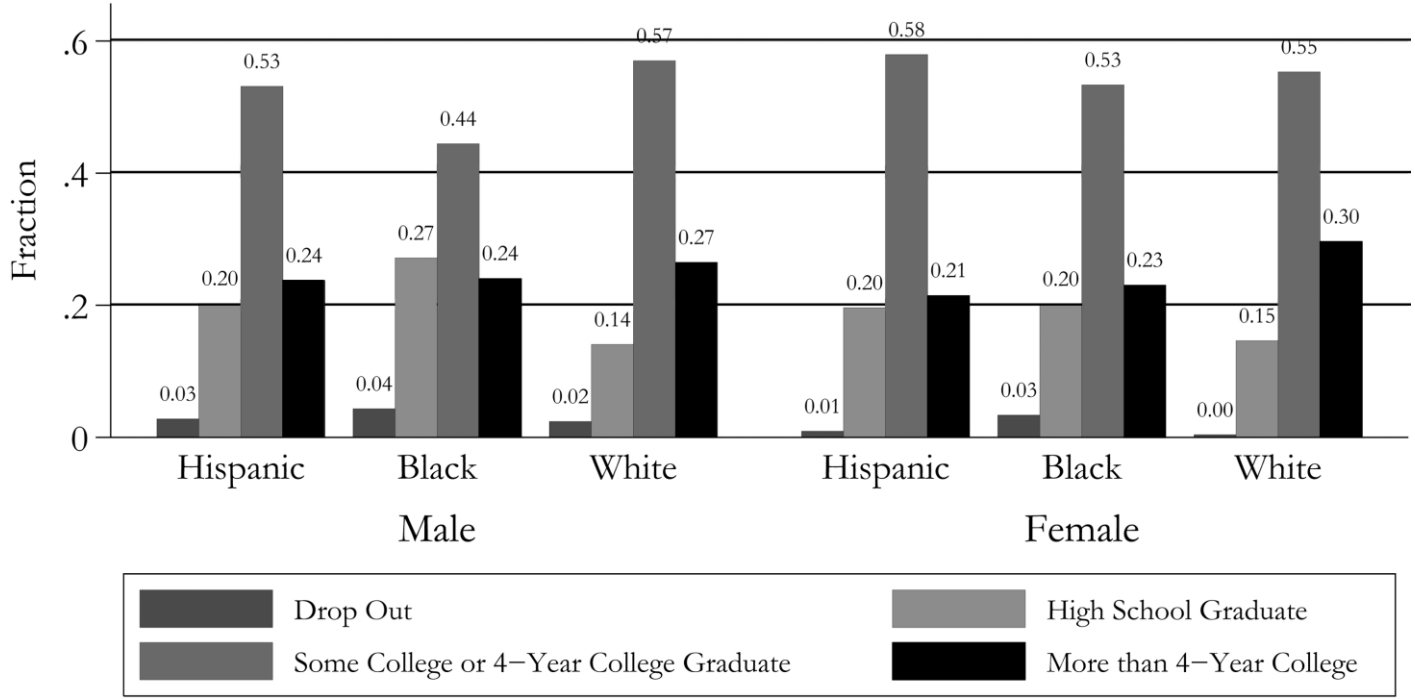

FIGURE 10.—Child's own expected educational level at age 10 by race and sex for CNLSY79 males and females. Each bar represents the number of people who report falling in a particular educational level cell divided by the total number of people in their race and sex group.

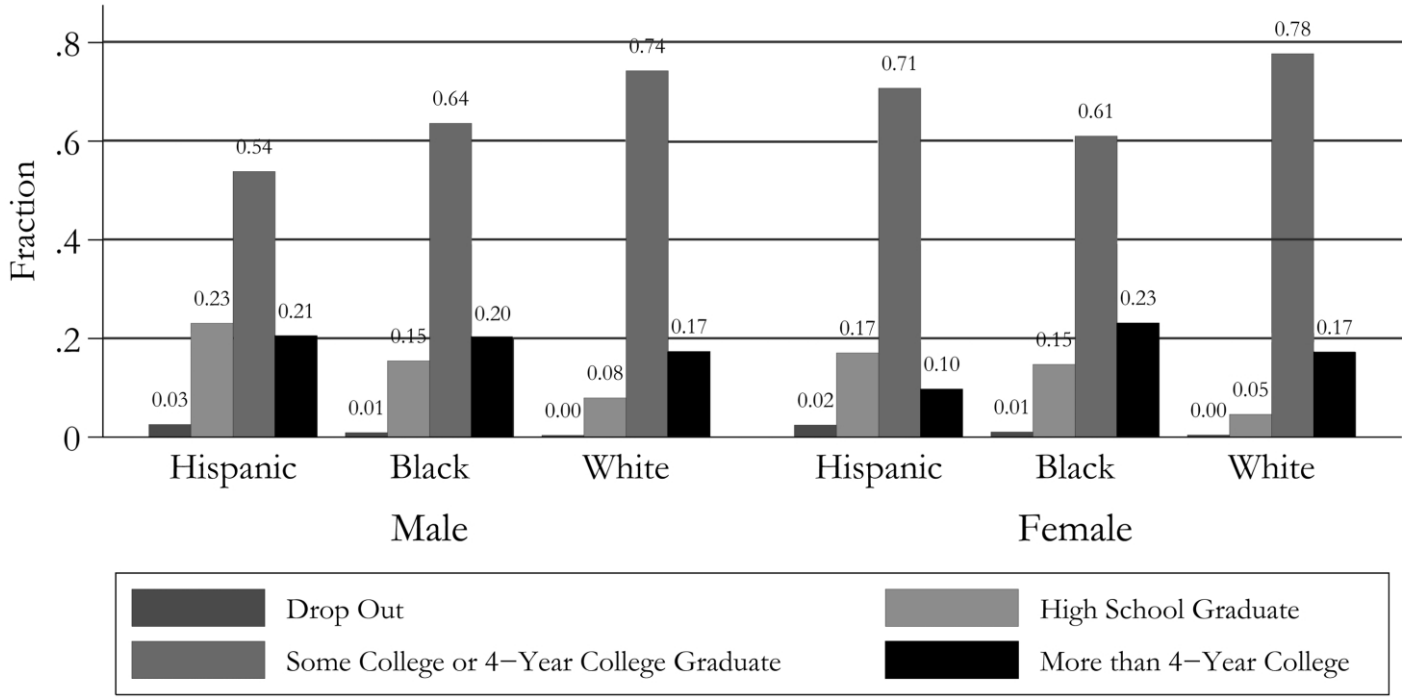

FIGURE 11.—Mother's expected educational level for the child at age 6 by race and sex for CNLSY79 males and females. Each bar represents the number of people who report falling in a particular educational level divided by the total number of people in their race and sex group.

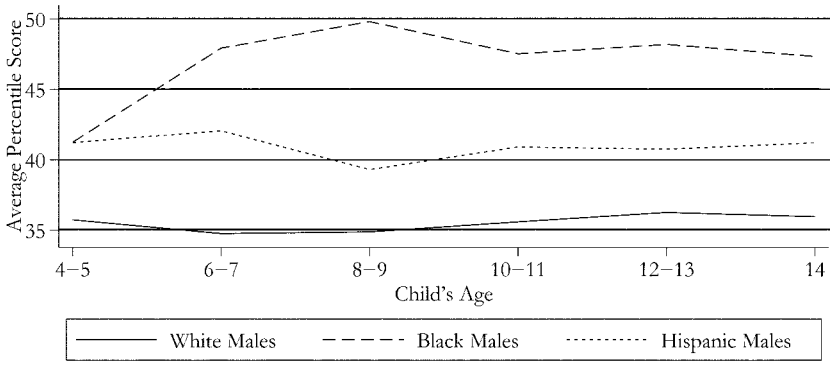

FIGURE 12.—Average percentile antisocial behavior score by race and age group for CNLSY79 males.

rud, and Urzua,[82] documents that noncognitive skills—motivation, self-control, time preference, and social skills—are important in explaining socioeconomic success.[83]

The CNLSY has life cycle measures of noncognitive skills. Mothers are asked age-specific questions about the antisocial behavior of their children such as aggressiveness or violent behavior, cheating or lying, disobedience, peer conflicts, and social withdrawal. The answers to these questions are grouped in different indices.[84] Figure 12 shows that there are important racial

[82] James Heckman, Jora Stixrud, & Sergio Urzua, Evidence on the Importance of Cognitive and Noncognitive Skills on Social and Economic Outcomes (unpublished manuscript, Univ. Chicago 2004).

[83] Some of the best evidence for the importance of noncognitive skills in the labor market is from the General Education Development (GED) program. This program examines high school dropouts to certify that they are equivalent to high school graduates. In its own terms, the GED program is successful. James J. Heckman & Yona Rubinstein, The Importance of Noncognitive Skills: Lessons from the GED Testing Program, 91 Am. Econ. Rev. 145 (2001), shows that GED recipients and ordinary high school graduates who do not go on to college have the same distribution of AFQT scores (the test graphed in Figure 1). Yet GED recipients earn the wages of high school dropouts with the same number of years of completed schooling. They are more likely to quit their jobs, engage in fighting or petty crime, or be discharged from the military than are high school graduates who do not go on to college or other high school dropouts. Intelligence alone is not sufficient for socioeconomic success. Minority-white gaps in noncognitive skills open up early and widen over the life cycle.

[84] The children's mothers were asked 28 age-specific questions about frequency, range, and type of specific behavior problems that children ages 4 and over may have exhibited in the previous 3 months. Factor analysis was used to determine six clusters of questions. The responses for each cluster were then dichotomized and summed to produce a raw score. The percentile score was then calculated separately for each sex at each age from the raw score. A higher percentile score indicated a higher incidence of problems. The antisocial behavior index we use in this paper consists of measures of cheating and telling lies, bullying and cruelty to others, not feeling sorry for misbehaving, breaking things deliberately (if age is less than 12), disobedience at school (if age is greater than 5), and trouble getting along with teachers (if age is greater than 5).

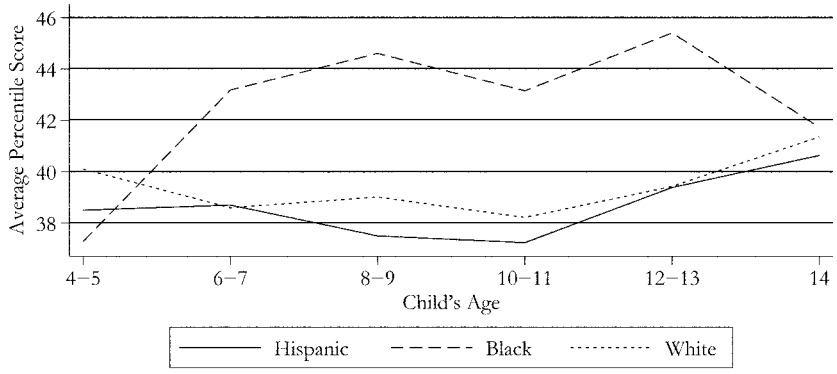

FIGURE 13.—Adjusted percentile antisocial behavior score by race and age group for CNLSY79 males.

and ethnic gaps in the antisocial behavior index that emerge in early childhood. The higher the score, the worse the behavior. By ages 5–6, the average black is roughly 10 percentile points above the average white in the distribution of this score.[85] The results shown in Figure 13, where we adjust the gaps by permanent family income, mother's education, and age-corrected AFQT and home scores, also show large reductions.[86]

Section II documents that minority and white children face substantial differences in family and home environments while growing up. The evidence presented in this section shows that these early environmental differences account (in a correlational sense) for most of the minority-white gap in noncognitive skills, as measured in the CNLSY.

Carneiro and Heckman[87] document that noncognitive skills are more malleable than cognitive skills and are more easily shaped by interventions. More motivated children achieve more and have higher measured achievement test scores than less motivated children of the same ability. The largest effects of interventions in childhood and adolescence are on noncognitive skills that promote learning and integration into the larger society. Improvements in these skills produce better labor market outcomes and less engagement in criminal activities and other risky behavior. Promotion of noncognitive skill is an avenue for policy that warrants much greater attention.

[85] In our Web appendix (*supra* note 9), we show that these differences are statistically strong. Once we control for family and home environments, gaps in most behavioral indices disappear.

[86] See our Web appendix, tables 2A and 2B, for the effect of adjusting for other environmental characteristics on the antisocial behavior score (*id.*).

[87] Carneiro & Heckman, *supra* note 21.

VII. Summary and Conclusion

This paper discusses the sources of wage gaps between minorities and whites. For all minorities but black males, adjusting for the ability that minorities bring to the market eliminates wage gaps. The major source of economic disparity by race and ethnicity in U.S. labor markets is in endowments, not in payments to endowments.

This evidence suggests that strengthened civil rights and affirmative action policies targeted at the labor market are unlikely to have much effect on racial and ethnic wage gaps, except possibly for those specifically targeted toward black males.[88] Policies that foster endowments have much greater promise. On the other hand, this paper does not provide any empirical evidence on whether the existing edifice of civil rights and affirmative action legislation should be abolished. All of our evidence on wages is for an environment in which affirmative action laws and regulations are in place.

Minority deficits in cognitive and noncognitive skills emerge early and then widen. Unequal schooling, neighborhoods, and peers may account for this differential growth in skills, but the main story in the data is not about growth rates but rather about the size of early deficits. Hispanic children start with cognitive and noncognitive deficits similar to those of black children. They also grow up in similarly disadvantaged environments and are likely to attend schools of similar quality. Hispanics complete much less schooling than blacks. Nevertheless, the ability growth by years of schooling is much higher for Hispanics than for blacks. By the time they reach adulthood, Hispanics have significantly higher test scores than do blacks. Conditional on test scores, there is no evidence of an important Hispanic-white wage gap. Our analysis of the Hispanic data illuminates the traditional study of black-white differences and casts doubt on many conventional explanations of these differences since they do not apply to Hispanics, who also suffer from many of the same disadvantages. The failure of the Hispanic-white gap to widen with schooling or age casts doubt on the claim that poor schools and bad neighborhoods are the reasons for the slow growth rate of black test scores. Deficits in noncognitive skills can be explained (in a statistical sense) by adverse early environments; deficits in cognitive skills are less easily eliminated by the same factors.

We have reexamined the Neal-Johnson[89] analysis that endowments acquired before people enter the labor market explain most of the minority-majority wage gap. Neal and Johnson use an ability test taken in the teenage years as a measure of endowment unaffected by discrimination. They omit schooling in adjusting for racial and ethnic wage gaps, arguing that schooling

[88] However, even for black males, a substantial fraction of the racial wage gap can be attributed to differences in skill.

[89] Neal & Johnson, *supra* note 6.

choices are potentially contaminated by expectations of labor market discrimination. Yet they do not adjust their measure of ability by the schooling attained at the date of the test, which would be the appropriate correction if their argument were correct.

Adjusting wage gaps by both completed schooling and the schooling-adjusted test widens the wage gaps for all groups. This adjustment effect is especially strong for blacks. Nonetheless, half of the black-white male wage gap is still explained by the adjusted score. At issue is how much of the majority-minority difference in schooling at the date of the test is due to expectations of labor market discrimination and how much is due to adverse early environments. While this paper does not settle this question definitively, test score gaps emerge early and are more plausibly linked to adverse early environments. The lion's share of the ability gaps at the date of the test emerge very early, before children can have clear expectations about their labor market prospects.

The analysis of Sackett, Hardison, and Cullen[90] and the emergence of test score gaps in young children cast serious doubt on the importance of stereotype threats in accounting for poorer black test scores. It is implausible that young minority test takers have the social consciousness assumed in the stereotype literature. If true, black skills are understated by the tests, and the market return to ability should be different for blacks than for whites. We find no evidence of such an effect.

Gaps in test scores of the magnitude found in recent studies were found in the earliest tests developed at the beginning of the twentieth century, before the results of testing were disseminated and a stereotype threat could have been "in the air." The recent emphasis on the stereotype threat as a basis for black-white test score differences ignores the evidence that tests are predictive of schooling attainment and market wages. It diverts attention away from the emergence of important skill gaps at early ages, which should be a target of public policy.

Effective social policy designed to eliminate racial and ethnic inequality for most minorities should focus on eliminating skill gaps, not on discrimination in the workplace of the early twenty-first century. Interventions targeted at adults are much less effective and do not compensate for early deficits. Early interventions aimed at young children hold much greater promise than strengthened legal activism in the workplace.

Bibliography

Altonji, Joseph, and Blank, Rebecca. "Gender and Race in the Labor Market." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David E. Card, pp. 3C:3143–4259. New York: Elsevier Science Press, 1999.

[90] Sackett, Hardison, & Cullen, *supra* note 18.

Armor, David J. *Maximizing Intelligence.* New Brunswick, N.J.: Transaction Publishers, 2003.

Ben-Porath, Yoram. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (1967): 352–65.

Black, Sandra E., and Sufi, Amir. "Who Goes to College? Differential Enrollment by Race and Family Background." Working Paper No. w9310. Cambridge, Mass.: National Bureau of Economic Research, 2002.

Bornholz, Robert, and Heckman, James J. "Measuring Disparate Impacts and Extending Disparate-Impact Doctrine to Organ Transplantation." *Perspectives in Biology and Medicine* 48 (2005): S95–S122.

Bowles, Samuel; Gintis, Herbert; and Osborne, Melissa. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39 (2001): 1137–76.

Bureau of Labor Statistics. *NLS Handbook 2001*. Washington, D.C.: U.S. Department of Labor, 2001.

Cameron, Stephen V., and Heckman, James J. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109 (2001): 455–99.

Campbell, Frances; Ramey, Craig; Pungello, Elizabeth; Sparling, Joseph; and Miller-Johnson, Shari. "Early Childhood Education: Young Adult Outcomes from the Abecedarian Project." *Applied Developmental Science* 6 (2002): 42–57.

Carneiro, Pedro, and Heckman, James J. "Human Capital Policy." In *Inequality in America: What Role for Human Capital Policies?* edited by James J. Heckman and Alan B. Krueger, pp. 77–240. Cambridge, Mass.: MIT Press, 2003.

Cunha, Flavio, and Heckman, James J. "The Technology of Skill Formation." Paper presented at the Society of Economic Dynamics and Control, Florence, 2004, and at the American Economic Association annual meeting, San Diego, 2004.

Cunha, Flavio; Heckman, James J.; Lochner, Lance; and Masterov, Dimitriy V. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of Education Economics,* edited by Finis Welch and Eric Hanushek. New York: North Holland, forthcoming 2005.

Currie, Janet, and Thomas, Duncan. "School Quality and the Longer-Term Effects of Head Start." *Journal of Human Resources* 35 (2000): 755–74.

Donohue, John J., and Heckman, James J. "Continuous versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks." *Journal of Economic Literature* 29 (1991): 1603–43.

Duncan, Greg, and Brooks-Gunn, Jeanne. *Consequences of Growing up Poor.* New York: Russell Sage, 1997.

Ferguson, Ronald. "What Doesn't Meet the Eye: Understanding and Addressing Racial Disparities in High Achieving Suburban Schools." Special

Edition, Policy Issues Report. Naperville, Ill.: North Central Regional Educational Laboratory, 2002.

Ferguson, Ronald. "Why America's Black-White School Achievement Gap Persists." Unpublished manuscript. Cambridge, Mass.: Harvard University, 2002.

Fryer, Roland, and Levitt, Steven. "Understanding the Black-White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 86 (2004): 447–64.

Hansen, Karsten; Heckman, James J.; and Mullen, Kathleen. "The Effect of Schooling and Ability on Achievement Test Scores." *Journal of Econometrics* 121 (2004): 39–98.

Heckman, James J. "Policies to Foster Human Capital." *Research in Economics* 54 (2000): 3–56.

Heckman, James J.; Larenas, Maria Isabel; and Urzua, Sergio. "Accounting for the Effect of Schooling and Abilities in the Analysis of Racial and Ethnic Disparities in Achievement Test Scores." Working paper. Chicago: University of Chicago, Department of Economics, 2004.

Heckman, James J.; Lochner, Lance; and Taber, Christopher. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics* 1 (1998): 1–58.

Heckman, James J., and Rubinstein, Yona. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91 (2001): 145–49.

Heckman, James J.; Stixrud, Jora; and Urzua, Sergio. "Evidence on the Importance of Cognitive and Noncognitive Skills on Social and Economic Outcomes." Unpublished manuscript. Chicago: University of Chicago, 2004.

Heckman, James J., and Todd, Petra. "Understanding the Contribution of Legislation, Social Activism, Markets and Choice to the Economic Progress of African Americans in the Twentieth Century." Unpublished manuscript. Chicago: American Bar Foundation, 2001.

Herrnstein, Richard, and Murray, Charles. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press, 1994.

Jencks, Christopher, and Phillips, Meredith. *The Black-White Test Score Gap*. Washington, D.C.: Brookings Institution, 1998.

Murray, Charles. "The Secular Increase in IQ and Longitudinal Changes in the Magnitude of the Black-White Difference: Evidence from the NLSY." Paper presented at the Behavior Genetics Association meeting, Vancouver, 1999.

Neal, Derek. "The Measured Black-White Wage Gap among Women Is Too Small." *Journal of Political Economy* 112 (2004): S1–S28.

Neal, Derek, and Johnson, William. "The Role of Premarket Factors in Black-

White Wage Differences." *Journal of Political Economy* 104 (1996): 869–95.

Sackett, Paul; Hardison, Chaitra; and Cullen, Michael. "On Interpreting Stereotype Threat as Accounting for African American–White Differences in Cognitive Tests." *American Psychologist* 59 (2004): 7–13.

Steele, Claude, and Aronson, Joshua. "Stereotype Threat and the Test Performance of Academically Successful African Americans." In *The Black-White Test Score Gap*, edited by Christopher Jencks and Meredith Phillips, pp. 401–27. Washington, D.C.: Brookings Institution, 1998.

Urzua, Sergio. "The Educational White-Black Gap: Evidence on Years of Schooling." Unpublished manuscript. Chicago: University of Chicago, Department of Economics, 2003.