The Dissertation Committee for Christopher Laurence Bartlett
certifies that this is the approved version of the following dissertation:

Essays on the Links Between Education, Ability, and Income

Committee:

Stephen J. Trejo, Supervisor

Stephen Donald

Robert A. Duke

Daniel S. Hamermesh

Gerald S. Oettinger

Essays on the Links Between Education, Ability, and Income

by

Christopher Laurence Bartlett, B.B.A.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2007

For Terri

ACKNOWLEDGMENTS

The help of others has made this work possible. I owe special thanks to Stephen Trejo, Daniel Hamermesh, and Gerald Oettinger for their helpful comments and suggestions. I would also like to thank Stephen Donald and Bob Duke. My wife, Terri Bartlett, has provided vital intellectual, editorial, and emotional support. My parents, Larry and Debbie Bartlett, instilled in me a sense of curiosity and the desire to finish what I start. I am indebted to Charles Hurd for providing me the opportunity to pursue an education. Finally, I must thank the Lord for giving me the ability and creating the circumstances that have put me in the position I am in today.

Essays on the Links Between Education, Ability, and Income

Publication No._________

Christopher Laurence Bartlett
The University of Texas at Austin, 2007

Supervisor: Stephen J. Trejo

This dissertation focuses on various aspects of the relationship between education and ability, paying particular attention to how the relationship may differ between races. Education and ability are important determinants of income, but the role they play differs by race. I focus on three questions. First, how does parental education affect child education? Second, why do black and white individuals of the same ability level choose different levels of education, and what impact does this have on wages? Finally, what effects do preferential college admission policies have on wages of college graduates?

The goal of the first chapter is to determine whether a mother's education level directly affects her child's education level, or whether the correlation in education levels can be attributed to other factors common to both the mother and child. I find that controlling for ability of both parent and child, additional maternal education has little effect on child education. For mothers with low levels of education I can not rule out a direct link between education levels, although I do produce estimates that are smaller than much of the current literature.

In the second chapter I develop a model that explains why, conditional on ability, blacks get more education then whites and yet have lower average wages. I then test the model empirically using two separate data sets. The results in general support the model.

The final chapter examines the effects of preferential college admission policies on labor market outcomes. There is little research linking school admission policies to labor market outcomes. Treating a school's selectivity as a signal of ability for its graduates, I show that rational employers will in effect undo preferential admission. This is because preferential admission both lowers the expectation and increases the variance of ability for preferred group. Using the Baccalaureate and Beyond Longitudinal Study of 1993 I construct a measure of preferential admission and empirically test the model. I find that for black graduates there is a significant decrease in wages due to affirmative action. This loss is similar in magnitude to the gain in wages related to school quality.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

The Extent of the Intergenerational Spillover of Education

1. INTRODUCTION

A strong correlation exists between the economic outcomes of parents and their children. Although this fact is widely known, the mechanism creating the correlation is poorly understood. The strong correlation between the education levels of parents and their children accounts for much of the persistence of economic outcomes, but this does not address the real question. Education level and economic outcome ultimately measure the same thing.

This paper examines the inter-generational transmission of education. The goal is to determine whether a parent's education level directly affects their children's education level or whether the correlation in education levels can be attributed to other factors common to both the parents and the children. A causal link between parent and child education would likely occur either because education makes people better parents or it reduces the cost of the child's education. Alternatively, the correlation could be due to similarities in parent and child intelligence, values, or culture.

The answer to this question has many policy implications. If there is a causal link in education, then a relatively short lived policy can have a lasting effect on a group's economic standing. Alternatively, if the correlation between two generation's education levels is due merely to other correlates such as ability, then education might be a poor tool for permanently increasing the economic outcome of a particular group.

I approach the problem in two ways. First, I utilize education obtained by parents late in life and argue that most of the correlation must be due to selection. Second, I take a more traditional approach but attempt to accurately control for both child and parent ability level. I find that much but not all of the correlation between parent and child educational outcomes can be attributed to selection. Additional maternal education has little effect on the child's education for children of mothers with high education levels. For mothers with low levels of education, I cannot rule out a direct link between mother and child education levels, although I do produce estimates that are smaller than those in much of the current literature.

1.1. Previous Work

Because of the importance of the topic, many researchers have studied the inter-generational transmission of human capital. The main hurdle facing researchers in this field is that education choices are endogenous and therefore are correlated with unobservable variables relating to environment and ability. To solve this problem, previous research has taken one of three main approaches: using twins to control for genetic factors, using adoptions under the assumption that children are assigned randomly, or exploiting exogenous changes in minimum education laws to instrument for a parental education. No consensus has been reached as to the causal extent of the correlation between parent and child education. Even researchers using similar method have come to very different conclusions.

Behrman and Rosenzweig (2002) look at the children of identical twins. This allows them to difference out the genetic factors affecting a child's educational outcome. They find that after controlling for genetic factors, father's education maintains a significant positive correlation with a child's education, while this correlation for mothers becomes negative and marginally significant. They argue that increased schooling leads to less time spent at home for women and is therefore detrimental to the child's education. Their argument that both maternal and paternal education directly affect the child's education, but in opposite directions, is weak. First, it relies heavily on the assumption that twins with different levels of education have the different levels of education because of some exogenous factor. This assumption seems questionable. Furthermore, it is a leap to claim that after genetic factors are controlled for, parental education must be driving the variation in child education. There are many other environmental factors shared by parent and child that will alter education decisions. This point will be explored more thoroughly shortly.

Plug (2002) and Sacerdote (2000) both look at adopted children and argue that the children are assigned randomly with no correlation in genetic factors between parents and adopted children. Although their empirical methods are similar, they find conflicting results. Sacerdote finds that most of the educational correlation can be attributed to environmental factors. Plug finds that the correlation between mother and child education can be attributed almost entirely to a correlation in ability, while the father-child correlation is due to both ability and environment. It is questionable whether adoptions are assigned at random. The studies also suffer from the problem that education is not the only environmental factor affecting children, and some of the other environmental factors are likely shared with the parents.

One note I have briefly mentioned in my discussion of both the adoption and the twin studies is that they address a slightly different question than the one I intend to examine. By controlling for genetic factors, the researchers are only able to determine whether parent-child education correlations are due to ability or environment. I am interested in whether an increase in parental education has a direct effect on child education or whether the correlation is merely due to selection. The difference is subtle but important. The family studies that attribute the education correlation to ability fall into what I refer to as selection, but the studies that find the correlation is due to environment do not necessarily fall into the category I refer to as causation. A parent's education is only one of the environmental factors affecting a child's educational outcome, there are many environmental characteristics shared by both parents and children that will affect education levels. From a policy perspective, knowing whether the correlation is due to causation or selection is more important than whether it is due to ability or environment. Knowing that the correlation is due to environment does not give information about the extent of inter-generational spillovers of education and does not provide insight into the efficacy of education policy as a tool. Parental education as well as other variables shared by parents and children are included in the category of environment.

Several papers use exogenous shifts in minimum schooling laws to instrument for parent's endogenous education choice. These papers also come to contradictory conclusions. Chevalier (2002) shows that the change in

British compulsory attendance laws had a small but significant impact on the children of the people affected by those laws. He argues that there is not a real endogenaity problem. Black, Devereux, and Salvanes (2003) perform a similar analysis looking at the change in Norwegian compulsory attendance laws and find that there were no spillovers. Oreopoulos, Page, and Stevens (2003) use U.S. Census data, aggregated to the state level, to preform a similar analysis. They find that the probability a child drops out decreases by between two and seven percent with each additional year of education required of the parents. One problem with using compulsory education laws is that they are only relevant to the very bottom of the educational distribution. The change in the laws will have no impact on the choices made by most students. It is also questionable whether the changes in laws are truly driving these results. A trend of increasing education levels for the entire population could plausibly cause an increase in the education levels of that population's youth as well as provide impetus for increasing minimum education requirements.[1]

Finally, Rosenzweig and Wolpin (1994) does not fit nicely into any of the other three categories. In this paper, Rosenzweig an Wolpin look at test scores of siblings, and uses difference in their mother's education at the time of their birth to control for education. He finds that an increase in a mother's education from ten to twelve years increases the child's test scores by about five percent. This approach most closely resembles the one I take, but there are many significant differences. First, there are likely birth order differences in children that are not accounted for by Rosenzweig and Wolpin. Also, a mother may be more mature and better able to raise

[1] Black, Devereux, and Salvanes (2003) are able to eliminate some time trend problems because their sample consists of municipalities that changed their laws at different times. Using this approach these researchers find little support for the intergenerational spillovers of education.

a child when she is older (with twelve years of education) than she was when she had her first child (with ten years of education). She will also be a more experienced parent when the second child comes along. These problems are exacerbated by the fact that Rosenzweig uses young mothers. I instead will look at changes in a mother's education as they relate to her child's education *after* the child has made the relevant education decision.

2. DATA

The data come primarily from the National Longitudinal Survey of Youth 1979 Children and Young Adults.[2] The Survey follows the biological children of women in the NLSY79 from 1984 to 2002. I examine three primary indictors of educational attainment for children. The first is whether a child drops out of school. This variable must be constructed. An individual is given a zero for this variable if they report their maximum education level as twelve or higher. They are given a one for this variable if their maximum reported education level is less than twelve and they either report this level twice, or the grade is less than or equal to their age minus nine,[3] where age is taken from the survey year in which they report their maximum grade. Thus, a twenty year old child reporting a maximum grade of eleven will be considered a dropout regardless of whether the maximum is reported twice or not. This gives a sample of 2,105 children and their mothers. Restricting the sample to those individuals with all of the relevant variables recorded reduces the sample to 1,929. This is the sample

[2] The unemployment rate is taken from the Bureau of Labor Statistics.

[3] Different rules for determining when an individual is considered a dropout were examined. In particular, two more restrictive samples were considered. In the first I required a maximum reported grade of less than or equal to age minus ten. To meet this and not have dropped out and individual would have to be three or four years behind their expected grade. I also looked at a sample requiring that the maximum grade be less than twelve, reported twice, *and* less than or equal to age minus nine. Neither of these rules had a significant impact on the results in the following sections.

used in section 5. The sample used in section 4 is slightly larger because individuals need not have a valid ability level reported.

The second variable I use to describe a child's educational attainment is an indicator for whether or not a child ever attends college. Individuals are given a zero for this variable if their maximum reported education level is less than thirteen and they either report this level twice, or it is more than two years lower than the appropriate grade for the individuals age,[4] where age is taken from the survey year in which they report their maximum grade. They are given a one if they have a maximum reported grade of thirteen or greater. This sample is smaller than the first because individuals reporting a maximum grade of twelve only once and who are under twenty-one when they do so are lost. It consists of $1,532$ children and their mothers. Once the sample is restricted to those individuals with all of the relevant variables, the sample contains $1,393$ individuals.

The final variable of interest is the highest grade completed. Individuals in this sample are given their maximum reported grade level if the maximum grade level is reported twice or is less than or equal to their age minus nine. This is the smallest sample because many individuals do not meet these criteria even though I can determine whether they dropped out or went to college. Restricting the sample to individuals with valid entries for all of the relevant variables give a sample size of $1,083$. The individuals included in the other two samples but not in this sample are those whom I observed finishing high school and then stopped reporting their highest grade completed or started college but stopped reporting their education level before they finished.

[4]The actual rule is the grade must be less than or equal to the individual's age minus nine.

TABLE 1

Averages and percents for key variables. HGC is higest grade completed, HGO is highest grade observed. The percentage of mothers who obtained more education after their children left school is recorded as MR. Mom's AAB represents the mothers age at birth. PPVT is a general measure of ability.

	N	MR	Age at HGC	HGC	Mom HGC	Mom's AAB	Black	Male	PPVT
HGC Identifiable									
Dropout	616	12%	17.5	9.5	10.9	19.8	48%	59%	79
High School	389	10%	19.5	12	11.5	18.9	46%	52%%	86
Some College	78	10%	22.5	13.7	11.7	17.7	46%	29%	91
HGC Unidentifiable	N	MR	Age exit Sample	HGO	Mom HGO	Mom's AAB	Black	Male	PPVT
At least H.S.	536	9%	18.4	12	12.2	20.8	39%	44%	90
At least some college	310	13%	21	14.1	12.5	19.6	36%	41%	95
Full Samples									
Dropout Sample	1929	11%	18.9	11.6	11.7	19.8	43%	50%	87
College Sample	1393	11%	19	11.5	11.5	19.4	45%	52%	85
HGC Sample	1083	11%	18.6	10.7	11.16	19.3	47%	54%	82

Table 1 gives the means and percentages of some key variables, breaking down each sample by the child's education level. The first three rows describe children for whom I can identify a highest grade completed. The fourth and fifth rows describe children who leave the sample before they finish their education, but for whom I can make some inference. When examining whether a child drops out, the first 5 rows of Table 1 can all be used. When looking at a child's college decision, the fourth row cannot be used because it is unclear whether a child in this category continued on to college or not. When looking at the highest grade completed, only the first three rows can be used. The bottom three rows give the variables for the samples used in section five. The table points out that mothers with higher levels of education tend to have children with higher levels of education. It also points to out that children achieving higher levels of education also have higher ability, as measured by the Peabody Picture Vocabulary Test-Revised (PPVT). This test is given to young children (over three) and should therefore not be influenced by the quality of the child's education. This test will be discussed more fully in section 5. These results suggest that mother's education, child's education, and ability are all highly correlated.

It is clear that the samples used are not representative of the population. They tend to be skewed toward individuals with less education because they are more likely to be observed finishing their education by 2002. This problem is exacerbated by the fact that the NLSY79 over-samples blacks, Hispanics, and economically disadvantaged non-minorities. This may cause serious issues because the sample will tend to select both parents and children with lower levels of education, and the extent to which the children are not representative of the larger population will vary with the mothers level of education. My sample of children of mothers with low levels of education looks similar to the original NLSY79 sample of such children as these mothers are likely to have children younger. On the other hand, the children of mothers with higher levels of education in my sample are more likely to be those children that did poorly in school relative to other children of highly educated mothers. This is because mothers with higher levels of education tend to have kids later in life, making it less likely that their children are observed finishing school. Thus, the problem is mitigated in the first two samples because an individual not observed finishing their education can still be in the college or dropout sample. The unrepresentative sample does have the benefit of representing the individuals with whom policy makers tend to be most concerned when discussing the benefits of increased education.

3. EDUCATION, ABILITY, AND CHOICE

In this section I introduce a simple model describing the education choice faced by children. A child chooses the optimal education by weighing the benefits of higher wages against the cost of extra education. If a parent's education level has a direct effect on the child's decision it is through reducing the cost of education[5] for the child. Thus, the child

[5] An argument can be made for a mother's education having an effect on a child's wages. This could happen if parents with more education have better connections and thus get their children better jobs. Although this is conceivable, the effects should be small compared to the effects through the decrease in cost. This is especially true if parental income is controlled for.

If I do allow parental education to directly affect wages, the results of this section hold with the additional assumption $\frac{\partial C}{\partial S_P} < \frac{\partial^2 W}{\partial S \partial S_P}$. The reduction in cost due to more parental income is smaller than the change in the marginal wage rate (with respect to education) due to more parental income.

solves the following problem:

$$\max_S U(S) = W(S,a) - C\left(a, a_P, S_p, Y\right) S \tag{1}$$

$$\frac{\partial W(S,a)}{\partial S} = C\left(a, a_P, S_p, Y\right) \tag{2}$$

where S and S_P are respectively the education levels of the child and the parent, a and a_P are the innate ability levels of the child and the parent, and Y is the parent's income. The function $W(S,a)$ represents the present discounted value of lifetime income, and $C(\cdot)$ is the cost of an additional unit of education. I assume that C is decreasing in a as more able people can acquire schooling with less investment. I also assume that C is decreasing in a_P. This assumption implies that a child with more able parents will face a lower cost of education even after controlling for the child's ability and the parent's education and income. The idea here is that more able parents will find it easier to help their children acquire education. The solution to the child's problem is given by equation (2). Deriving the effects of a change in parental education on a child's education is a straightforward application of the implicit function theorem. Specifically,

$$\frac{\partial S}{\partial S_P} = \frac{\partial C}{\partial S_P} \Delta^{-1} \tag{3}$$

where $\Delta = \frac{\partial^2 W}{(\partial S)^2} < 0$. Similarly, the effects of a change in ability are given by

$$\frac{\partial S}{\partial a} = \left(-\frac{\partial^2 W}{\partial S \partial a} + \frac{\partial C}{\partial a} \right) \Delta^{-1} \tag{4}$$

Equation (3) gives the direct effect of a change in parental education on the child's education if all other variables remain static. In application of the model to data, the controls for a and a_P are typically few and poor. By applying equation (4) to the parents, one can see that parental education is a function of parental ability $S_P = S_P(a_P)$ with $\frac{\partial S_P}{\partial a_P} > 0$. It is also reasonable to assume that the child's ability is correlated with parental ability. Thus I will assume $a = a\left(a_P\right)$ with $\frac{\partial a}{\partial a_P} > 0$.

$$\frac{\partial S}{\partial S_P} = \left(-\frac{\partial^2 W}{\partial S \partial a} \frac{\partial a}{\partial a_P} \frac{\partial a_P}{\partial S_P} + \frac{\partial C}{\partial S_P} + \frac{\partial C}{\partial a_P} \frac{\partial a_P}{\partial S_P} + \frac{\partial C}{\partial a} \frac{\partial a}{\partial a_P} \frac{\partial a_P}{\partial S_P} \right) \Delta^{-1} \tag{5}$$

Equation (5) gives the effect of a change in parental education on the child's education if a and a_P are not controlled for. Note that I assume that $S_P(a_P)$ has an inverse $a_P\left(S_P\right)$. This equation has multiple parts. The first term describes the effects on income through the increase in the child's ability associated with the increase in parental education. The second portion is the direct increase due to a reduction in cost of education. The third gives the decrease in cost due to the parent's higher ability level, and the last gives the decrease in cost due to the increase in the child's ability. It is clear that the partial effects must be larger if parental and child ability are not accounted for. Equations (3) and (5) together show that when ability is not accounted for, the returns to parental education will be overestimated. I want to estimate what is represented by equation (3), but without properly accounting for mother and child ability levels the estimate will be represented by equation (5). I have not made assumptions on $\frac{\partial C}{\partial S_P}$ but it is typically assumed to be negative. Even if it is positive, as claimed by Behrman and Rosenzweig (2002), the result holds.

4. AN ARGUMENT FOR SELECTION

In this section I will argue that what we observe as the transmission of education is primarily due to selection. To do this I will exploit the fact that a significant portion of parents, approximately 11%, return to school after their child has either graduated or dropped out of high school. I use the differences in correlation between the educational outcomes of children and their parents in different groups to argue that much of this correlation must be due to selection.

To begin, I propose a thought experiment. Imagine two groups of parents, each with their education level set at the birth of their child. In one group all of the parents have a high school diploma, and in the other they all have a college diploma. After controlling for all of the relevant variables, the econometrician finds that the children of the college group have a seventy percent probability of attending college, while the children of the high school group have only a fifty percent chance of going to college. There are two possible explanation for the observed differences. Either the groups differ in unobservable ways that influence a child's education, or the parents in the college group are able to reduce the cost of education for their children *because* they went to college.[6]

Now imagine a third group that I will call the switchers. The parents in this group have a high school diploma when their children are eighteen, but go to college after their child has made his decision about college. If the children of the switchers are found to have a fifty percent probability of going to college, then we are no closer to answering the original question. It could be that the high school group and the college group differ in unobservable ways, with the switching group resembling the high school group, or it could be the direct effect of having parents who have gone to college before the child goes through school. Alternatively, if the children of the switchers have a seventy percent chance of going to college, the se-

[6] Allowing for a combination of the two possibilities does not change the conclusions drawn from the example. I will use the polar cases for clarity.

lection argument is the only one that makes sense. It must be that the high school and college groups differ in unobservable ways and the switching group resembles the college group. To argue that causation also exists would require that the switching parents differed from both other groups, and that had they gone to college before their children were born, the kids would be even more likely than those of the college group to go to college.[7] While this is not unimaginable, it seems unlikely.

In reality we know that the correlation between a parent and her child's educational outcome is due to both a direct causal effect as well as selection. The true goal is to measure the portion of the correlation caused by each of these. I include the above example merely to clarify the first approach taken in doing this. The argument can be formalized as follows. The reduced form equation describing education for child j in family i is given by

$$S_{ij} = \delta_0 + \delta_1 S_i^M + \delta_2 S_i^F + \gamma_1 a_i^M + \gamma_2 a_i^F + \gamma_3 a_{ij} + \Gamma h_i + \beta X_i + \varepsilon_i \quad (6)$$

where S represents education obtained before the child makes the relevant decision, a represents ability, and h represents other heritable characteristics. The vector X represents non-heritable child characteristics and ε represents a child specific error. The superscripts M and F indicate the given variable for the mother and father, respectively. The inclusion of h in the model elucidates one drawback of the twin and adoption studies. The variable h can represent environmental and cultural factors shared by both parents and children. This may differ between twin mothers and therefore will not be removed by the twin studies. Notice that S^M is the education

[7] It could also be argued that going to college has a direct negative effect on one's children's education. This too is not unimaginable but seems unlikely.

obtained by the mother before the child decides to go to college or drop out. I will also assume that some mothers choose to obtain more education after their child has finished schooling denoted by dS^M. Thus, a mother's total education is given by $TS^M = S^M + dS^M$. Ability measures a^M and a as well as the heritable characteristics h_i are not observable. To compensate for this I construct linear projections of them on the observables. These are given by equation (7) through (9).[8]

$$a^M = d_0 + d_1 TS^M + d_2 X \tag{7}$$

$$a = t_0 + t_1 TS^M + t_2 X \tag{8}$$

$$h = g_0 + g_1 TS^M + g_2 X \tag{9}$$

I assume that the variables in X are child specific and not related to the unobservables so that $d_2 = t_2 = g_2 = 0$. In the following empirical work I do not use data on fathers, but father's education and ability can be described by the following assortative mating equations.

$$S_i^F = r_1 S_i^M + r_2 a_i^M + e_i \tag{10}$$

$$a_i^F = b_1 S_i^M + b_2 a_i^M + v_i \tag{11}$$

Positive r's and b's imply positive assortative mating on both education and ability.

The parameter of particular interest is δ_1. The difficulty in estimating

[8] Note that this implicitly assumes that $cov\left(a^M, S^M\right) = cov\left(a^M, dS^M\right)$: ability plays the same role in determining education both before and after the child leaves school. This is a strong assumption that may not hold. As long as $cov\left(a^M, dS^M\right)$ is between zero and $cov\left(a^M, S^M\right)$, the results that follow will hold but they will not be as strong. The procedure will remove some but not all of the bias.

Similar statements can be made about $cov\left(a, S^M\right) = cov\left(a, dS^M\right)$ and $cov\left(h, S^M\right) = cov\left(h, dS^M\right)$.

this parameter comes from all of the unobservables included in equation (6). To remove these I take advantage of the fact that the education obtained by mothers after their child leaves school is not directly in equation (6) but is related to the unobservables. Thus substituting equations (7) through (11) into equation (6) gives:

$$S = [\delta_1 + \delta_2 r_1 + \gamma_2 b_1]\, S^M + [\gamma_3 t_1 + \Gamma g_1 + (\gamma_1 + \delta_2 r_2 + \gamma_2 b_2)\, d_1]\, TS^M + X\beta + \nu \tag{12}$$

where ν contains the new constant plus the combined error terms. Defining the coefficient on a mother's total education as $\lambda = \gamma_3 t_1 + \Gamma g_1 + (\gamma_1 + \delta_2 r_2 + \gamma_2 b_2)\, d_1$ and rearranging gives

$$S = [\delta_1 + \delta_2 r_1 + \gamma_2 b_1 + \lambda]\, S^M + (\lambda)\, dS^M + X\beta + \nu \tag{13}$$

The coefficient of interest is δ_1, the direct effect of a mother's education on her child's education. As equation (13) shows, this is not simply the estimated coefficient on a mothers education level. By estimating equation (13) and differencing the estimated coefficients on a mother's education acquired before and after her child leaves school, I can find $\widehat{\delta_1} = \delta_1 + \delta_2 r_1 + \gamma_2 b_1$. This gives an estimate of the effects of a mother's schooling on her child's schooling that is still biased upward, but with a smaller bias. It is actually giving the effects of increasing a mother's education including those associated with her finding a better husband. This effect may be of interest to policy makers if they are considering improving the education of a select number of individuals, but will be less important if they are discussing raising the education of the country or a large region. In the absence of assortative mating $\widehat{\delta_1} = \delta_1$ and the estimate is the true effect of an increase in mother's education on child education.

To apply this idea to the data I separate mothers into two groups, each of which I analyze identically but separately. First, I look only at mothers with twelve or more years of school when their children leave high school, either by graduating or dropping out. All three groups from the original example are represented by these parents. I regress both the number of years of college education before and after the child left school on an indicator variable for whether or not the child went to college. This is done using both a linear probability model and probit estimation. The results are similar and coefficients on the variables of interest are reported in Table 2.[9] Because of the straightforward interpretation I will refer to the results of the linear regression.

It can be seen that each additional year of college attained before the child leaves school is correlated with a 5% increase in the likelihood the child goes to college, while each additional year of college attained by the parents after the child leaves high school is correlated with a 3.4%[10] increase in this likelihood. Both of these results are statistically significant. This indicates that approximately seventy percent of the educational correlation between parents and children is due to selection. In fact, this estimate is a lower bound. This is because selection would account for seventy percent of the correlation only if the group of parents who obtain more education after their children leave high school do not differ in unobservable ways

[9] Along with the independent variables given in the second column of Table 2, the independent variable included in each regression are: percentage of time the father lived in the household, an indicator for gender, indicators for race, an indicator for wether the child had kids before they were eighteen, the mothers age at birth linearly and squared, interactions of race and the percentage of time the father lived at home, interactions of race and kids, and an interaction of kids and percentage of time the father lived at home.

[10] The 5% estimate is represented in the reduced form theory as $\delta_1 + \delta_2 r_1 + \gamma_2 b_1 + \lambda$, where λ =3.4%

TABLE 2

Coefficients and Standard Errors From Educational Outcome Regression by Group

		Linear Regressions		Probit Regressions	
		College	Dropout	College	Dropout
Mothers With H.S. Diplomas When Child Leaves School	Years College Before Child Left School	0.050	−0.015	0.142	−0.061
		(0.012)	(0.008)	(0.033)	(0.032)
	Years College After Child Left School	0.034	−0.005	0.096	−0.022
		(0.015)	(0.011)	(0.043)	(0.043)
	N	841	1267	841	1267
Mothers That Never Attend College	H.S. to H.S. Indicator	0.149	−0.167	0.563	−0.525
		(0.029)	(.026)	(0.108)	(0.083)
	No H.S. to H.S. Indicator	0.085	−.068	0.245	−0.238
		(0.124)	(0.127)	(0.455)	(0.418)
	N	881	1227	881	1227

from the parents who obtain the same level of education before their children leave school. In reality these unobservable characteristics probably fall somewhere in between the parents who obtain college years before their children leave school and those who never go to college.

Unfortunately, this is the only regression that gives any information. Using the same group but looking at the probability a child drops out of school gives an insignificant coefficient on the years of college obtained after the child leaves school. As pointed out earlier, this does not provide any information about the cause of the correlation. It could be that δ_1 is large and significant, or alternatively the assumptions $cov\left(a^M, S^M\right) = cov\left(a^M, dS^M\right), cov\left(a, S^M\right) = cov\left(a, dS^M\right)$ and $cov\left(h, S^M\right) = cov\left(h, dS^M\right)$ may fail.

A similar analysis is performed on a second set of parents. These are parents who never attend college. The independent variables used for this set are indicators for whether a parent had a high school diploma before the child left school or received a diploma after the child left school, with parents never receiving a diploma as the base group. These results are mixed. The point estimates indicate that a large portion of the correlation can be attributed to selection. Unfortunately, the standard errors for the switching indicator variable are large, indicating that the children of the switching parents may not differ significantly from those with parents that never receive a high school diploma.

These results highlight the fact that the people who return to school after their children leave high school differ in unobservable ways from those that do not. This can also be seen by looking at the education levels of the grandparents of the students. Table 3 gives the average education of students' grandparent and oldest aunt or uncle broken down by the

TABLE 3
Comparison of family education by parental group.

Mother's Education Before/After Child stops	Average Highest Grade Completed by Grandmother	Grandfather	Oldest Aunt	N
College/College	10.57	10.36	12.71	490
H.S./College	10.20	10.17	12.52	61
H.S./H.S.	10.07	9.79	11.77	681
No H.S./H.S.	8.50	10.42	11.90	14
No H.S./No H.S.	8.27	7.91	10.84	659

parent's education level. The grandparents of the two groups of children whose parents change education level do not resemble those of any of the groups of parents not changing education level.

5. THE IMPORTANCE OF ABILITY

The major concern when using standard OLS or probit regressions of a mother's education on that of her child has to do with unobserved ability. Ability poses two major problems. First, there is a correlation between the unobserved ability levels of mother and child and a link between education and ability. The concern is that mothers with high ability levels get more education and also have children with high ability levels who in turn are more likely to achieve high levels of education. Second, it is likely that mothers with higher ability levels are better able to raise their children regardless of their education. Twin and adoption studies of human capital transmission are designed specifically to remove the effects of ability. One particularly beneficial aspect of the NLSY is that it contains several measures of ability for both mothers and children. When ability is omitted from the education regression, the coefficient on mother's education will be biased upward because it will pick up some of the correlation in ability. By including ability in the regressions much of the bias can be removed. The drawback to this approach is the same as that of the twin and adoption studies. If a child's educational outcome looks like their parent's outcome because they share similar environmental factors other than education, these will not be removed. The approach only allows for the separation of ability from environment, not education from ability and non-education environment. Thus the coefficient provides an upper bound for the effects of education.[11]

In this section I regress different educational outcomes of children on their mother's education and ability, as well as a measure of the child's ability. Including a measure of the child's ability removes the correlation in mother and child education through a link in their ability. By including a measure of mother's ability I can account for the fact that mothers with higher ability might be better parents. Two measures of child ability are utilized. The first is the Peabody Picture Vocabulary Test-Revised (PPVT). This test is given to children three years old or older. The second is the Peabody Individual Achievement Test (PIAT) Math score which is taken after the child is age five. Both tests are standardized[12] by age. I measure the mother's ability using the 1980 version of the Armed Forces Qualifying Test. These are also standardized by age.

[11] The Children of the NLSY are given the Home Observation for Measurement of Environment Short Form. This is a test designed to measure the quality of the child's home environment. I plan on incorporating this into the model in the future. It is not included in this version because the test measurement error will be correlated with the mothers education and I do not have a good instrument for the test.

[12] The population mean is 100 with a standard deviation of 15.

5.1. Empirical Methodology

The equation I wish to estimate is:

$$edu = X\beta_0 + \beta_1 \,(mother\ edu) + \gamma_1 a^c + \gamma_2 a^m + \varepsilon \tag{14}$$

where *edu* represents the child's educational outcome of interest, a^c and a^m represent ability of child and mother, and X is a matrix containing other explanatory variables. Both the PPVT and PIAT Math provide noisy measures of ability. Under the assumptions I make on the measurement error, all β_0 and β_1 will be inconsistent. The estimate of γ_1 will also suffer from attenuation bias. This will cause the estimate $\widehat{\beta_1}$ to likely overestimate the true β_1. To account for this I use the multiple indicator solution.[13] This involves instrumenting one of the ability measures with the other. First, each measure of ability is written to be a linear combination of true ability and a random error. Thus:

$$PPVT = \delta_0 + \delta_1 a + r_1 \tag{15}$$

$$PIAT = \rho_0 + \rho_1 a + r_2 \tag{16}$$

It is apparent that the error in the measurement of ability is necessarily correlated with the measurement itself. I assume that the measurement errors r_1 and r_2 are not correlated with each other, true ability a, X, or mother's education.[14] Solving equation 15 for a and plugging this into

[13] See Wooldridge, pp.105-107.

[14] The strongest assumption listed here is that $cov(r_1, r_2) = 0$. If some children are just good test takers this would be violated. I argue that it is not. The tests used are given at different times and use different testing procedures. Further, I repeated the full analysis using two other measures of ability with no significant change in the results.

equation 14 gives

$$edu = X\beta_0 + \beta_1 \left(mother\ edu\right) + \frac{\gamma_1 \delta_0}{\delta_1} + \frac{\gamma_1}{\delta_1} PPVT + \gamma_2 a_m + \left(\varepsilon - \frac{\gamma_1}{\delta_1} r_1\right) \tag{17}$$

The correlation between $PPVT$ and r_1 causes problems that can be removed using IV estimation. Because that $PPVT$ and $PIAT$ are only correlated through their dependence on a, $PIAT$ makes a good instrument.[15] Thus I estimate

$$PPVT = X\lambda_0 + \lambda_3 \left(mother\ edu\right) + \lambda_2 PIAT + u \tag{18}$$

and use the predicted values of $PPVT$ in place of the actual values. A benefit of using this method over leaving ability in the error term and trying to instrument mother's education is that the multiple indicator solution allows for any variables in X to be correlated with ability. This is important because it is highly likely that a mother's ability and other demographic variables are in fact correlated with ability. As mentioned earlier, the drawback is that it does not control for unobserved non-education environmental factors shared by parents and their children.

5.2. Results

I examine three measures of education; the probability a child's begins college, the probability a child drops out of high school, and the highest grade completed (HGC) by the child for those children I can observe leaving school. The first two measures are dichotomous and are estimated using both OLS and Probit specifications. Table 4 reports the estimated coef-

[15] The reverse is also true, $PPVT$ makes a good instrument for $PIAT$. I performed the analysis both ways with no significant difference in the results.

ficients on mother's education. Each outcome is examined for the entire sample of children and for two subsamples. The first subsample consists of children whose mothers have at least a high school diploma, and the second subsample consists of children whose mothers have at most a high school diploma. I split the sample this way because it seems likely that the effects of an extra year of college education may differ greatly from the effects of an extra year of high school education. Only individuals with valid values for all explanatory variables are used, giving the sample sizes reported in the last column. Because of the straightforward interpretation I will discuss the linear estimates for the two dichotomous outcomes.

The first column of table 4 reports the coefficients on the mother's education variable when controls for ability are not included. These estimates are consistent with much of the previous literature. Using a sample of over one quarter of a million individuals, Black, Devereux, and Salvanes (2003) find that each additional year of education obtained by the mother leads to an estimated 0.15 year increase in the educational attainment of the child. They also report almost identical answers when they estimate the returns to mothers education for mothers with less than eleven years of education. Chevalier (2002) finds that an extra year of mother's education increases the probability a child completes post-compulsory schooling by about 4%, and Oreopoulos, Page, and Stevens (2003) find that it decreases the probability of dropping out by 2% to 7%. My OLS estimates of the probability a child drops out match nicely with both of these estimates. Finally, Sacerdote (2000) finds that the extra year of mother's education increases the probability a child attends college by 3% to 7%.[16]

One of the most interesting results reported in table 4 is the progression

[16] All of the reported estimates are from the non-instrumented results in the specified papers. They should be compared to my OLS results.

of coefficients as one moves to the right of the table. When a measure of ability is included, the coefficient on the mother's education falls by approximately 60% in the $p(College)$ model, 50% in the $p\,(Dropout)$ model, and 20% in the highest grade completed model. When ability is included in the estimation but is not instrumented, the coefficient on ability should be biased downward because of measurement error. This leads to a likely upward bias in the coefficient on mother's education. This can be seen when looking at the coefficient estimates from the instrumented models. In every specification for every sample, the OLS estimate without ability controls is larger then the OLS with controls, which is larger than the IV estimate.[17]

Table 4 indicates that there are decreasing marginal returns (in terms of a child's education) to a mother's education. The coefficient on the mother's education remains significant and is larger for mothers with at most a high school diploma than it is for mothers with at least a high school diploma. For the latter group, the coefficients from the IV regression are always insignificant, and the coefficients from the OLS estimates with ability controls are insignificant in all except the regression on whether a child drops out, and even this is insignificant in the probit specification. Once ability is accounted for, a mother's post high school education has little impact on her child's education, but her pre-college education does. This relationship holds for all measures of child educational attainment whether ability is instrumented or not. It also holds in both the linear and probit estimates.

[17] It should be noted that the IV estimates reported here do not suffer from the typical problem of large standard errors because they are not the coefficient on the variable being estimated.

		Linear Probability Model			Probit Model			
		OLS	OLS	2SLS IV	Probit	Probit	Probit IV	N
	Ability Controls	No	Yes	Yes	No	Yes	Yes	
p(College)	Full Sample	0.0437 (0.005)	0.0167 (0.006)	0.0105 (0.006)	0.1520 (0.018)	0.0737 (0.022)	0.0459 (0.024)	1,393
	Mothers with Edu≥12	0.0420 (0.011)	0.0148 (0.012)	0.0119 (0.012)	0.1189 (0.033)	0.0427 (0.036)	0.0320 (0.038)	778
	Mothers with Edu ≤12	0.0742 (0.014)	0.0443 (0.014)	0.0371 (0.015)	0.2974 (0.054)	0.2112 (0.058)	0.1887 (0.061)	842
p(Dropout)	Full Sample	−0.0436 (0.005)	−0.0231 (0.005)	−0.0171 (0.006)	−0.1398 (0.015)	−0.0784 (0.019)	−0.0585 (0.020)	1,929
	Mothers with Edu≥12	−0.0144 (0.008)	−0.0026 (0.008)	−0.0007 (0.008)	−0.0592 (0.032)	−0.0104 (0.034)	−0.0060 (0.036)	1,181
	Mothers with Edu ≤12	−0.0720 (0.013)	−0.0490 (0.012)	−0.0445 (0.013)	−0.2176 (0.039)	−0.1471 (0.042)	−0.1297 (0.043)	1,184
HGC	Full Sample	0.1537 (0.024)	0.1263 (0.029)	0.1026 (0.030)	…	…	…	1,083
	Mothers with Edu≥12	0.0330 (0.040)	0.0008 (0.042)	−0.0020 (0.042)	…	…	…	530
	Mothers with Edu ≤12	0.1892 (0.046)	0.1489 (0.047)	0.1439 (0.047)	…	…	…	665

TABLE 4

Coefficients and standard errors from regression of education outcome. Ability controls included are the mother's AFQT 1980 and the child's PPVT score. Where IV is indicated, PPVT is instrumented with PIAT. Other included variables are percentage of time the father lived in the household, an indicator for gender, indicators for race, and indicator for whether the child had kids before they were eighteen, the mothers age at birth linearly and squared, interactions of race and the percentage of time the father lived at home, interactions of race and kids, and an interaction of kids and percentage of time the father lived at home.

Decreasing marginal returns to a mother's education make sense. Children with mothers having high levels of education are more likely to have access to tutors, library books, and other factors that make their mother's knowledge less important in their education. Children whose mothers have low levels of education are more likely to rely on their mothers for help because they may not have access to tutors and other factors that may decrease the cost of education. Decreasing marginal returns to mother's education also suggest that unobserved non-education environmental factors are important. It could be that some families put a high value on education, which would be picked up in the mother's education coefficient. This might only show up in mothers with low levels of education if all of the families in the sample with higher education levels place roughly the same value on education.

The estimates in Table 4 also indicate, although only weakly, that a mother's education has the largest impact on whether or not her child drops out. An extra year of school obtained by a mother decreases her child's expected probability of dropping out by 1.7% and increases their probability of going to college by 1.1%. For mothers with at most a high school diploma, each additional year of school reduces her child's probability of dropping out by 4.5% and increases the child's probability of attending college by 3.7%. This relationship does not hold for mothers with higher levels of education, but estimated effects of more education on both the probability a child drops out and the probability a child goes to college are insignificant.

Taken as a whole, table 4 provides two major results. First, controlling for ability removes most of what is typically measured as inter-generational human capital transfer. Second, there is an inter-generational spillover of education for at low levels of maternal education,[18] but not at high levels. The table also indicates that the spillovers primarily occur at the bottom end of a child's educational outcome, although the support for this is weak. For mothers with at most a high school diploma, an extra year of education leads to an expected increase 3.7% in the probability their child goes to college, a 4.5% decrease in the probability their child drops out of school, and expected increase of 0.14 years of education for their child.

6. IMPLICATIONS AND CONCLUSION

It is important that policy makers know to what extent education carries over from generation to generation. Education levels differ systematically by group which leads to differing economic outcomes by group. In my sample 35% of blacks drop out of high school while only 23% go to college. Only 30% of whites drop out of high school while 29% go to college. Knowing the extent of education transmission gives an indication of the efficacy of education policy as a tool for changing group outcomes.

Comparing parental education obtained both before and after a child makes their relevant education decision, I find indications but not conclusive evidence that much of the correlation in education is due to selection. By controlling for ability of both mothers and children I find more conclusive evidence. The marginal effect of maternal education on child education is small for mothers with high levels of education. Alternatively, I cannot rule out a direct effect of maternal education on child education for mothers with low levels of education, although my estimates of the effect are small. I find that a mother's education has a small but significant effect

[18] These spillovers do apear to be smaller than those found in much of the current research.

on her child's education for mothers with low levels of education. Because this result could be due to a causal link or to non-education environmental factors, this group deserves further research. The transmission of education in families with low levels of education is of great interest because public policy is often targeted at precisely these families.

Chapter II

Ability, Education, and Dynamic Discrimination

1. INTRODUCTION

Conditioned on ability, as measured by the Armed Forces Qualifications Test (AFQT), blacks acquire more education than whites. This fact appears at odds with the wage gap between blacks and whites. Figure 1 shows that conditioned on ability, blacks have a higher level of education. If education is productive, then higher levels of education should lead to higher wages for any given level of ability, yet this relationship does not appear to hold when comparing groups of blacks to whites. In this paper I propose a model that can explain the black-white wage gap as well as account for the observed education-ability relationship. In fact, the same mechanism generates both relationships. I then test the model empirically using the National Longitudinal Survey of Youth 1979 (NLSY79) and the Baccalaureate and Beyond Longitudinal Study of 1993 (B&B). The results, in general, support the model.

The model I develop builds on two existing models of discrimination. It incorporates a dynamic model of statistical discrimination as in Oettinger (1996) and individual education decisions in an environment where education is observable but not a perfect predictor of productivity as in Lang and Manove (2004). In the model, worker quality is determined by both ability and education and is inferred correctly by employers, but productivity is also dependent on match quality which is not known by employers.

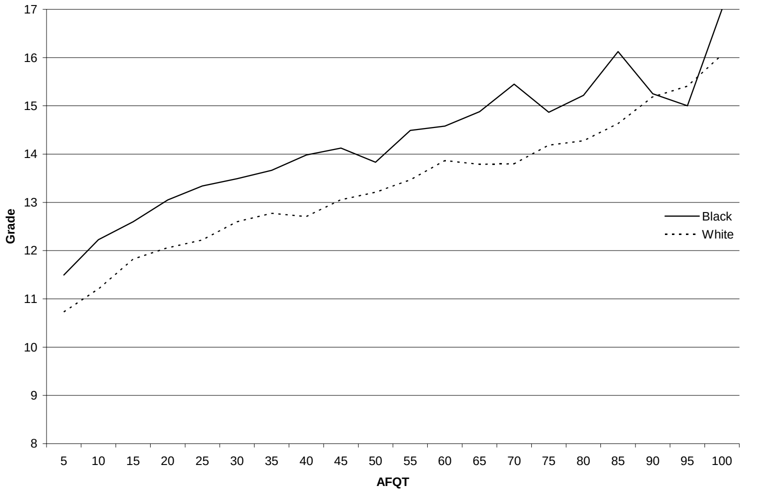

FIG. 1 Average highest grade completed for a given AFQT by race.

A signal exists as to true worker productivity, but the signal is less precise for blacks. This causes employers to weight inferred worker quality more heavily than the productivity signal for blacks, giving them an incentive to over invest in education. The same noisy signal of true productivity also increases the likelihood of poor job matches for blacks and thus slows their wage growth over time.

The model predicts that given the same ability level blacks will acquire more education than whites. This result is consistent with the findings of the static model Lang and Manove (2004).[19] Using a cross-sectional framework, their empirical results show that blacks are not compensated for this additional investment, and they attribute the lack of compensation to labor market discrimination. This contrasts much of the previous work

[19] The introduction of a dynamic setting does reduce the predicted size of the difference between black and white education levels.

on statistical discrimination, most notably Lundberg and Startz (1983), which attributes the wage gap to pre-market factors. The dynamic setting of my model allows for a wage gap to emerge over time. This gives rise to the possibility that, at any given job, blacks are fully compensated for their investment in education, yet the emergent wage differential may overtake the initial higher pay (conditional on education) received by blacks. Using cross-sectional data in this environment will incorrectly show that blacks are not fully compensated for their investment. Ignoring the dynamic aspect of the model leads to results that indicate individual employers pay blacks less than whites with comparable levels of education and ability. The real cause of the gap is the difficulty blacks face in distinguishing between good and bad job matches.

1.1. Previous work

Discrimination is the subject of a large body of research both within and outside the field of economics. Initial work in economics focused on tastes for different types of workers (Becker 1957), however this approach has fallen out of favor in recent years. It is difficult to construct a taste based model that generates long run differences in pay for blacks and whites. Most of the current research into discrimination can be classified as statistical discrimination. The basic premise of these models is that employers have imperfect information about worker productivity. Workers are paid their conditional expectation of productivity based on a noisy signal and observable characteristics. It is assumed that the variance of the black productivity signal is larger than that of whites, and thus employers put less weight on the signal for blacks.

As proposed by Aigner and Cain (1977), this model does not predict a difference in average wages between blacks and whites. Assuming ability has the same distribution for blacks and whites, the model only predicts that black wages will have a smaller variance. This is because employers will put more weight on the average productivity of a workers group as the signal for individuals in that group becomes noisier. Lundberg and Startz (1983) modify the basic model by adding an unobservable investment decision. Workers can invest in human capital which increases both their productivity and their noisy productivity signal. Because employers put less weight on the signal of blacks, there is less incentive for blacks to invest in human capital. This leads to blacks having less human capital and being less productive on average.

Although this model predicts a wage gap, it is not clear what the unobservable human capital investment is. Employers typically know an individual's level of education as well as their grades. It is also reasonable to assume that they know the quality of the most recent school attended.[20] In addition, most jobs require some sort of resume or written application on which an applicant is free to elaborate on their qualifications. Individuals who have made productivity enhancing investments have an incentive to inform the employer. There are unobservable factors that affect an individual's human capital and vary by race, but most of these are exogenous to the individual. An individual cannot choose the quality of their elementary school or the level of their parent's education.[21]

Oettinger (1996) develops a model similar to that of Aigner and Cain

[20] Most local employers have a general idea of the quality of the high schools in their area, and most firms hiring college graduates know the relative quality of the schools from which they hire. If they do not, this information is easy to obtain, so if school quality does matter it seems that firms would invest in obtaining ranking of schools.

[21] The Lundberg Startz (1983) model also has the unrealistic property that all of the individuals of a given race choose the same level of investment. This might be remedied with a simple change in the functional form, but it is not clear that all of the results will hold.

(1977) with the addition of multiple periods in which the agents are able to change jobs. In this model productivity is determined entirely by the quality of a job match and not by the ability of an individual. In the first period there is no average wage gap, but due to the larger variance in the productivity signal of the blacks, they find it harder to identify productive job matches in later periods. This creates a between race wage differential in the second period. The model produces results that match the data but is silent regarding the optimal level of investment in human capital for individuals. Knowing that they will have trouble identifying productive job matches, will blacks invest in more or less education?

In contrast to Lundberg and Startz (1983), Lang and Manove (2004) develop a model of statistical discrimination in which individuals can make an observable investment in human capital. This notion of investment is more appropriate when considering education and pre-labor market training. They find that blacks actually have a greater incentive to invest in education than do whites of the same ability level. Under these circumstances, measuring wages controlling for education but not ability will produce results that overestimate the wage differential. This is because, conditional on education, blacks have lower ability and will therefor earn lower wages. Similarly, conditioning on ability alone produces results that underestimate the wage differential. If their model is correct, controlling for both education and ability should eliminate the wage gap. Empirically they find that a wage gap does still exist, and they attribute it to labor market discrimination. My research attempts to explain why this wage gap remains.

In a departure from the majority of empirical literature, Neal and Johnson (1996) estimate a model that includes only age, AFQT, and race. They argue that other common explanatory variables are influenced by future expectations of labor market discrimination. By not including education or other choice variables they claim to measure the full effect of discrimination. This may be true, but is does not inform us as to whether different races are rewarded differently in the labor market.

This research attempts to answer two questions that have not yet been adequately addressed. First, how will individuals choose their level of human capital investment in a dynamic setting where blacks have more trouble identifying productive job matches? Second, why does a wage gap exist even after controlling for ability and education?

2. THE MODEL

Why does a between race wage gap exist even when blacks get more education than do whites of a given ability? I develop a model that attempts to explain both the higher level of observable human capitol investment of blacks and lower wages received by blacks. The model builds heavily on the models developed by Oettinger (1996) and Lang and Manove (2004). Specifically, it retains the education/ability relationship of Lang and Manove and the dynamic aspect of Oettinger.

The model combines aspects of statistical discrimination, sorting, and dynamic matching models. Worker productivity is a combination of worker quality and match quality. Firms can observe worker quality but only receive a noisy signal of match quality. Consistent with the statistical discrimination literature, I assume that the match quality signal for blacks is less informative than it is for whites. Match quality is revealed with the passage of time, and workers are allowed to change jobs. Controlling for education and ability, there will be no wage gap the first period but one will develop over time.[22] If ability is ignored, blacks will appear to face a wage gap that will only worsen over time. Controlling only for ability will show that whites face lower wages in the first period but make wage gains relative to blacks with the passage of time.

2.1. Environment and Assumptions

Consider a continuum of individuals, each with different exogenous pre-labor market ability[23] a. Individuals work for two periods and maximize net lifetime income. Before entering the labor market, workers must choose their level of education (s) based on their ability, expected lifetime income, and the cost of education. Productivity in each period is a function of education, ability, and match quality (ε). I assume that education and ability are complements. Neither workers nor firms know the value of the match quality until it has been tried for one period. Every individual receives a single job offer each period. In the second period the individual can remain at the first period job where true productivity is known, or move to a new job with uncertain productivity (but a known wage).

Formally, suppose that the productivity of worker i at the job offered in period t is given by

$$P_{it} = Q\left(s_i, a_i\right) + \varepsilon_{it} \tag{19}$$

where $Q\left(s, a\right)$ represents worker quality and is a function of education and ability. I assume $\frac{\partial^2 Q(s,a)}{\partial s \partial a} > 0$ so that education and ability are complementary inputs, and $\frac{\partial^2 Q(s,a)}{(\partial s)^2} < 0$ so that there are diminishing returns to

[22] This is consistent with Oettinger (1996), but rather than there being one average match productivity, there is now an average match productivity for every education-ability pair.

[23] Lang and Manove refer to this as innate ability, but it is more appropriate to think of it as the portion of ability that is exogenous to the individual and acquired before labor market entry. The distinction will be more fully discussed later.

education. Ability is distributed over some closed interval $a_i \in [\underline{a}, \overline{a}]$. No other assumptions are made on the distribution of ability. The quality of the employee and firm match is random and captured by $\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right)$, where ε_1 and ε_2 are independent and uncorrelated. Total productivity has both a deterministic and a random component.

Neither firms nor individuals know the match quality (ε_{it}) of a particular job offer until it has been tried, but there is an observable signal of true productivity

$$p_{it} = P_{it} + \nu_{it} \tag{20}$$

where $\nu_{it} \sim N(0, \sigma_\nu)$ represents a random error term in the productivity signal and is uncorrelated across individuals and time. Firms also observe an individual's investment in education (s_i) but do not directly observe ability (a_i). In period two, productivity is known to be P_{i1} for those workers who do not change jobs.

The remainder of this section will proceed as follows. I will first solve for the optimal wage offer for firms. I will then find the optimal level of investment for individuals. Using these, I will examine the model's predictions as to education, average wages, and discrimination. For the remainder of the paper I will suppress the i subscript to simplify notation.

2.2. Firms

Firms play no strategic role in my model. I assume that there is a competitive market and that firms pay workers their conditional expected productivity. To find the optimal wage contract, I assume that an individual's educational investment is described by the continuous function $s(a)$ which is known by firms. I also assume $s'(a) > 0$. After solving the workers problem, I will show that $s(a)$ does represent the solution. This solution stems from the complementary nature of s and a in the production function.

Firms see two basic types of workers, new employees and returning employees. In the first period all individuals are new workers. In the second period, those individuals that stay at their first period job are returning workers while those that change jobs are viewed as new workers once again. The true productivity of returning workers is known and they are paid their true productivity. Thus, $w_2 = P_1$ for those workers who remain at their first period job.

New workers are paid their expected productivity conditional on the observable productivity signal p_t and education s. Although firms do not directly observe ability, they know $s(a)$ and its inverse $a(s)$ and can therefor infer worker quality. I denote the firm inference of $Q(s,a)$ as $q(s,a(s))$, and the competitive equilibrium wage as $\widehat{w_t}(p_t,s)$. For these workers

$$\widehat{w_t}(p_t,s) = E[P_t|p,s] \tag{21}$$

$$E[P_t|p,s] = E[Q(s,a)|p,s] + E[\varepsilon|p,s] \tag{22}$$

Firms can infer $Q(s,a)$ so $E[Q(s,a)|p,s] = q(s,a(s))$. Combining equations (19) and (20) gives $p_t - Q(s,a) = \nu_t + \varepsilon_t$. Once again using the inference of worker quality $p_t + q(s,a(s)) = \nu_t + \varepsilon_t$, which is bivariate normally distributed with ε_t. This bivariate normality gives

$$E[\varepsilon|p_t - q(s,a(s)) = \nu_t + \varepsilon_t, s] = \lambda(p_t - q(s,a(s))) \tag{23}$$

where $\lambda = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\nu^2}$.

Thus the wage offer for new workers is

$$\widehat{w_t}(p_t, s) = q(s, a(s)) + \lambda(p_t - q(s, a(s))) \tag{24}$$

$$\widehat{w_t}(p_t, s) = \lambda p_t + (1 - \lambda) q(s, a(s)) \tag{25}$$

Examination of equation (25) shows that employers pay wages that are a weighted combination of the productivity signal and their inference of worker quality based on observable education. It is also apparent that as σ_ν^2 increases employers put more weight on their inference of quality and less on the productivity signal. The wage offer made by firms is:

$$\widehat{w_1} = \lambda p_1 + (1 - \lambda) q(s, a(s))$$

$$\widehat{w_2} = \begin{cases} P_1 & \text{if Returning} \\ \lambda p_2 + (1 - \lambda) q(s, a(s)) & \text{if New} \end{cases}$$

2.3. Workers

I turn now to the worker's problem. Each worker faces two decisions: their level of educational investment and whether to move to a new job or stay at their current job. I start by examining the movement decision. The worker makes this decision between periods one and two. At the time of the decision the worker knows true productivity at their current job and has a new job offer in hand. Thus, the worker only moves if $P_1 < \lambda p_2 + (1 - \lambda) q(s, a(s))$.

The worker must choose s before knowing $\varepsilon_1, \varepsilon_2, \nu_1$, or ν_2. His goal is to maximize lifetime income $V(a, s)$. The ex ante expected lifetime income can be written as

$$V(a, s) = E[w_1] + pr(\text{stay}) E[w_2|\text{stay}] + pr(\text{move}) E[w_2|\text{move}] - C(s) \tag{26}$$

which is equal to

$$
\begin{aligned}
E\left[w_1\right]+pr(P_1 &> \widehat{w_2}\left(p_2,s\right))E\left[w_2|P_1>\widehat{w_2}\left(p_2,s\right)\right] \\
+pr(P_1 &< \widehat{w_2}\left(p_2,s\right))E\left[w_2|P_1<\widehat{w_2}\left(p_2,s\right)\right]-C\left(s\right) \qquad (27)
\end{aligned}
$$

Where $C(s)$ represents the cost of investing in education. I will proceed by solving for each portion of (26) in terms of s, and then solve for the worker's optimal s. In the first period the worker can expect to receive $E\left[w_1|a,s\right]=E\left[\widehat{w}\left(p(a,s),a\right),s\right]$.

$$
\begin{aligned}
E\left[\widehat{w}\left(p(a,s),a\right),s\right] &= \lambda E\left[p_1|a,s\right]+\left(1-\lambda\right)E\left[q\left(s,a(s)\right)\right] \qquad (28)\\
E\left[\widehat{w}\left(p(a,s),a\right),s\right] &= \lambda E\left[Q|a,s\right]+\lambda E\left(\varepsilon+\nu\right)+\left(1-\lambda\right)E\left[q\left(s,a\left(s\right)\right)\right]
\end{aligned}
$$

so

$$E\left[w_1|a,s\right]=\lambda Q\left(s,a\right)+\left(1-\lambda\right)q\left(s,a\left(s\right)\right) \qquad (29)$$

Examination of equation (29) shows that employees expected wages in period one are a combination of their true quality and the firms inference about that quality. This allows workers to use education as a signal of quality as well as a true quality enhancer. Notice that as σ_ν^2 increases the signaling aspect of education becomes more important relative to its true impact on quality.

The second period expected wages vary depending on the movement decision. If the worker stays, expected wages are

$$E\left[w_2|\text{stay}\right]=E\left[P_1|P_1>\lambda p_2+\left(1-\lambda\right)q\left(s,a(s)\right)\right] \qquad (30)$$

which is equivalent to

$$E\left[w_2|\text{stay}\right] = Q\left(s,a\right) + E\left[\varepsilon_1|\varepsilon_1 - \lambda\left(\varepsilon_2+\nu_2\right) > \left(1-\lambda\right)\left(q\left(s,a\left(s\right)\right) - Q\left(s,a\right)\right)\right] \quad (31)$$

The bivariate normality of ε_1 and $(\varepsilon_1 + \lambda(\varepsilon_2 + \nu_2))$ implies[24] that

$$E\left[w_2|\text{stay}\right] = Q\left(s,a\right) + \frac{\sigma_\varepsilon}{\sqrt{1+\lambda}}\left[\frac{\phi\left(\beta\right)}{1-\Phi\left(\beta\right)}\right] \quad (32)$$

where ϕ represents the standard normal p.d.f., Φ represents the standard normal c.d.f., and $\beta = \frac{(1-\lambda)(q(s,a(s))-Q(s,a))}{\sigma_\varepsilon\sqrt{1+\lambda}}$. For the workers who move,

$$E\left[w_2|\text{move}\right] = E\left[\lambda p_2 + \left(1-\lambda\right)q\left(s,a(s)\right)|\lambda\left(\varepsilon_2+\nu_2\right) - \varepsilon_1 > \left(1-\lambda\right)\left(q\left(s,a\left(s\right)\right) - Q\left(s,a\right)\right)\right] \quad (33)$$

or equivalently

$$\begin{aligned} E\left[w_2|\text{move}\right] &= q\left(s,a(s)\right) + \lambda\left(Q\left(s,a\right) - q\left(s,a(s)\right)\right) \quad (34) \\ &\quad + E\left[\lambda\left(\varepsilon_2+\nu_2\right)|\lambda\left(\varepsilon_2+\nu_2\right) - \varepsilon_1 > \left(1-\lambda\right)\left(q\left(s,a\left(s\right)\right) - Q\left(s,a\right)\right)\right] \end{aligned}$$

Once again utilizing the bivariate normality of $(\lambda(\varepsilon_2 + \nu_2))$ and $(\lambda(\varepsilon_2 + \nu_2) - \varepsilon_1)$, I find

$$E\left[w_2|\text{move}\right] = q\left(s,a(s)\right) + \lambda\left(Q\left(s,a\right) - q\left(s,a(s)\right)\right) + \frac{\lambda\sigma_\varepsilon}{\sqrt{1+\lambda}}\left[\frac{\phi\left(\beta\right)}{1-\Phi\left(\beta\right)}\right] \quad (35)$$

The last portion of equation (26) needed before I can solve for s is the probability of moving. Recall that the worker only moves if $P_1 < \lambda p_2 +$

[24] See Green (2003) p. 781

$(1-\lambda)\, q\,(s, a(s))$ so

$$pr(\text{move}) = pr\left(P_1 < \lambda p_2 + (1-\lambda)\, q\,(s, a(s))\right) \tag{36}$$

or equivalently

$$pr(\text{move}) = pr\left(\varepsilon_1 - \lambda\,(\varepsilon_2 + \nu_2) > (1-\lambda)\left(q\,(s, a\,(s)) - Q\,(s, a)\right)\right) \tag{37}$$

or

$$pr(\text{move}) = \Phi\,(\beta) \tag{38}$$

Notice that in general the probability an individual moves is a function of s and a.

Combining (29), (32), (35), and (38) in (26) we can write the workers maximization problem as

$$\begin{aligned}
&\max_s \lambda Q\,(s, a) + (1-\lambda)\, q\,(s, a\,(s)) \\
&+ [1 - \Phi\,(\beta)] \left[Q\,(s, a) + \frac{\sigma_\varepsilon}{\sqrt{1+\lambda}} \left[\frac{\phi\,(\beta)}{1 - \Phi\,(\beta)}\right]\right] \\
&+ [\Phi\,(\beta)] \left[q\,(s, a(s)) + \lambda\left(Q\,(s, a) - q\,(s, a(s))\right) + \frac{\lambda \sigma_\varepsilon}{\sqrt{1+\lambda}} \left[\frac{\phi\,(\beta)}{1 - \Phi\,(\beta)}\right]\right] \\
&- C\,(s)
\end{aligned} \tag{39}$$

The solution to the workers optimal investment problem can be characterized by the differential equation

$$2Q_s + Q_a a'\,(s)\,(1-\lambda) \left[\frac{3}{2} - \frac{2\lambda}{\pi\,(1+\lambda)}\right] = C'\,(s) \tag{40}$$

Recalling that $a(s)$ is the inverse of $s\left(a\right)$, the solution can be written as

$$s'\left(a\right)=\frac{Q_a}{C'\left(s\right)-2Q_s}\left[\frac{3\pi-4\lambda-\lambda^2\left(3\pi-4\right)}{2\pi(1+\lambda)}\right] \tag{41}$$

Notice that this satisfies the initial assumption of $s'\left(a\right)>0$ as long as $C'\left(s\right)>2Q_s$. Recall also that $a_i\in\left[\underline{a},\overline{a}\right]$. These two facts pin down a solution to equation (41). To see this, note that for individuals with $a_i=\underline{a}$ there is no incentive to signal, and they will invest at the efficient level. This gives $s\left(\underline{a}\right)$ and creates a unique solution to (41).

In equation (40) the $2Q_s$ represents the direct effects of increased education on productivity. Additional investment in education both increases the expected productivity signal and increases firms inference of ability. The second left hand side term represents the inferred increased return to education when the individual does not follow $s(a)$.

2.4. Implications

According to the model, people with higher ability levels will acquire more education, as is evident in equation (41). To examine the model's predictions about discrimination, I will assume that the productivity signal for blacks is noisier than that of whites, as is usually done in the statistical discrimination literature. This implies that $\sigma_{uw}^2<\sigma_{ub}^2$ and $\lambda_w>\lambda_b$. Recall that the wage faced by new applicants is $\widehat{w_t}\left(p_t,s\right)=\lambda p_t+\left(1-\lambda\right)q\left(s,a(s)\right)$. Because firms can observe race, the wage offered to new black employees relative to that of new white employees will be weighted more heavily toward inferred quality and less heavily toward the productivity signal. Because education has a signalling component as well as a direct effect on inferred quality but has only a direct effect on the productivity signal, black individuals will have an incentive to invest more heavily in education than do white individuals.

To show formally that blacks get more education than whites I note that $s_w(\underline{a}) = s_b(\underline{a})$. As mentioned before, education has no signaling value at the lowest ability level so blacks and whites at $\underline{a}$ get the same level of education. It can also be seen from equation (41) that for every a such that $s_w(a) = s_b(a)$, $s'_w(a) < s'_b(a)$. Thus, for $a \in (\underline{a}, \overline{a})$ we have $s_w(a) < s_b(a)$. Empirically this implies that if we do not account for ability, we will overestimate the wage differential. This is because conditional only on education, blacks have lower ability and are paid less. Also, if ability is ignored, the return to education will appear lower for blacks than for whites. For a given quality blacks will have more education and less ability than whites.

To clarify I will provide a simple example.[25] Say $Q(a,s) = \min\{\ln a, \ln s\}$, $C(s) = \frac{s}{3}$, $\underline{a} = 1$ and $s(\underline{a}) = 1$. In this example the assumption of $s'(a) > 0$ does not hold for $a > s$, but for $a < s$ the model produces

$$s'(a) = \frac{1}{a}\left(\frac{3\pi - 4\lambda - \lambda^2(3\pi - 4)}{\frac{1}{3}2\pi(1+\lambda)}\right) \tag{42}$$

and

$$s(a) = \left(\frac{3\pi - 4\lambda - \lambda^2(3\pi - 4)}{\frac{1}{3}2\pi(1+\lambda)}\right)\ln a + 1 \tag{43}$$

Now assume that $\lambda_w = \frac{1}{2}$ and $\lambda_b = \frac{1}{4}$. Figure 2 shows a graph of the optimal level of investment in education as a function of ability for both whites and blacks.

The top line represents blacks optimal investment and the bottom represents that of whites. The dotted line gives efficient investment. In this

[25] This is a modified version of an example presented in Lang and Manove (2004).

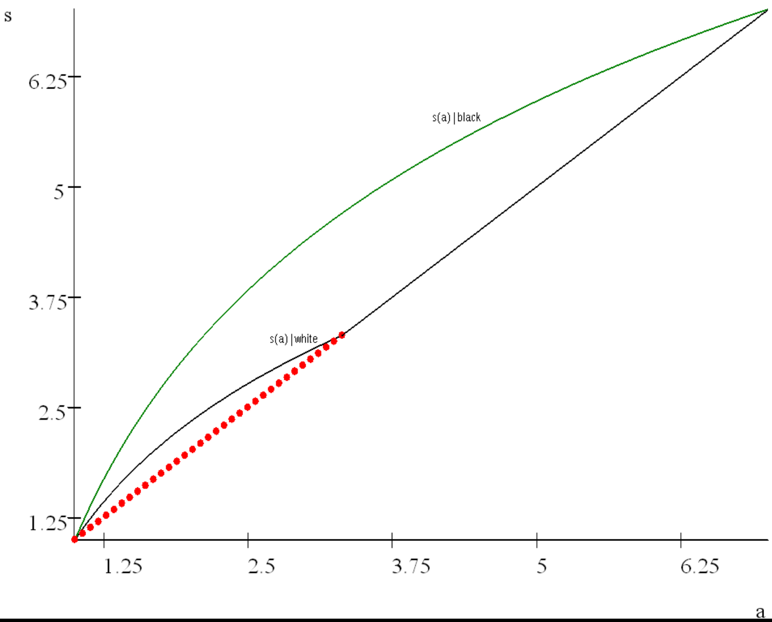

FIG. 2 Investment in education by race

example blacks get weakly more education at all levels of ability. There is also a range over which both groups over invest in education. In this region the cost of the nonproductive education is outweighed by the information it carries about a person's ability level.

The results up to this point parallel those of Lang and Manove (2004), but here they diverge. Lang and Manove use only a static model and thus do not explain any wage dynamics. Their empirical findings do not match the model and they attribute the difference to labor market discrimination. Specifically, they take cross-sectional data and control for education and ability. They find that blacks are not compensated for their extra investment in education. The dynamic aspect of my model explores why these results might arise in a cross-sectional study.

First look at average first period wages for a given ability education pair.

$$E\left[P_1\right] = \lambda E\left[Q(s,a)\right] + (1-\lambda)\, E\left[q\left(s, a\left(s\right)\right)\right] + \lambda E\left[\varepsilon + \nu\right] \tag{44}$$

In equilibrium firm inference about productivity is correct so $Q(s,a) = q\left(s, a\left(s\right)\right)$ and

$$E\left[P_1\right] = Q\left(s,a\right) \tag{45}$$

This finding is similar to Oettinger (1996) except that his average wages were taken over the entire group whereas these are for only those people of a given level of ability and education. One argument against this model is that if ability has the same distribution for blacks and whites, then the model predicts that blacks will have higher first period average wages. There are two reasons why we might not observe this even if the model is correct. First, it is hard to say how long a period is in reality. It is possible that workers learn match quality quickly and bad matches are terminated in a period shorter than the data collection time frame. If this is so the data will not pick up first period wages.

The second reason is related to the manner in which a is defined. I define a as exogenous pre-labor market ability rather than innate ability as Lang and Manove (2004). Under my definition there is no reason to believe that a has the same distribution in both populations. One would actually expected this a to be correlate with unobservable race related variables. All I require is that workers cannot choose to invest in a and that it does not change in the labor market. This definition of ability also fits well with the typical understanding of what AFQT measures. Studies have found that AFQT is a racially unbiased measure of ability (Wigdor and Green, 1991) but the distribution of AFQT is significantly lower for blacks than it is for whites. Figures 3 and 4 provide a comparison of the histograms of standardized AFQT for blacks and whites. Unobservables such as childhood environment and primary school quality likely affect AFQT but are still appropriately considered part of a in my model.

Now consider the second period wages. Stayers receive their first period productivity as their wage in period two. Once again drawing upon bivariate normality and the fact that $Q(s,a) = q(s, a(s))$, in equilibrium

$$E\left[P_1|P_1 > \lambda p_2 + (1-\lambda)\, q(s, a(s))\right] = Q(s,a) + \frac{\sigma_\varepsilon}{\sqrt{1+\lambda}}\left(\frac{2}{\sqrt{2\pi}}\right) \tag{46}$$

Similarly, I find average wages for movers in period two as

$$E\left[\lambda p_2 + (1-\lambda)\, q(s, a(s))\, |\lambda(\varepsilon_2 + \nu_2) - \varepsilon_1 > 0\right] = Q(s,a) + \frac{\lambda\sigma_\varepsilon}{\sqrt{1+\lambda}}\left(\frac{2}{\sqrt{2\pi}}\right) \tag{47}$$

Once again these findings are similar to Oettinger (1996) with the exception that now the average is conditional on ability and education. Comparing equations (46) and (47) with (45) reveals that all average wages increase in period 2, but the average wage of stayers increases by more than the average wage of movers. This should produce results showing that the measured returns to tenure and experience are both positive. Also, whites should have a larger return to experience and a smaller return to tenure.

Finally consider the average between period wage changes. For stayers the average change is

$$\begin{aligned} & E\left[P_1 - \lambda p_1 - (1-\lambda)\, q(s, a(s))\, |P_1 > \lambda p_2 + (1-\lambda)\, q(s, a(s))\right] \\ = \; & E\left[\varepsilon_1 - \lambda(\varepsilon_1 + \nu_1)\, |\varepsilon_1 - \lambda(\varepsilon_2 + \nu_2) > 0\right] = \frac{(1-\lambda)\,\sigma_\varepsilon}{\sqrt{1+\lambda}}\left(\frac{2}{\sqrt{2\pi}}\right) \end{aligned} \tag{48}$$

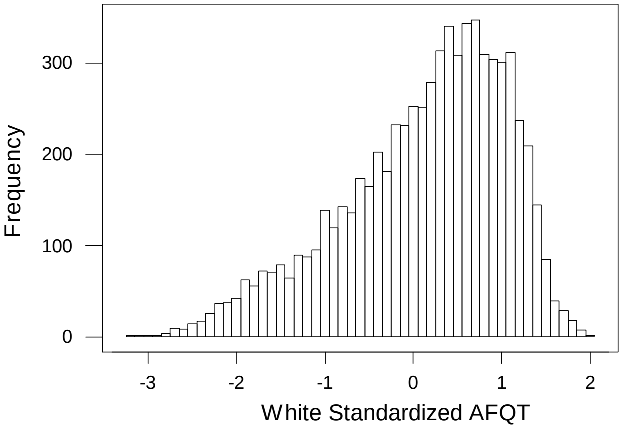

FIG. 3 Histogram of white AFQT scores standardized by age

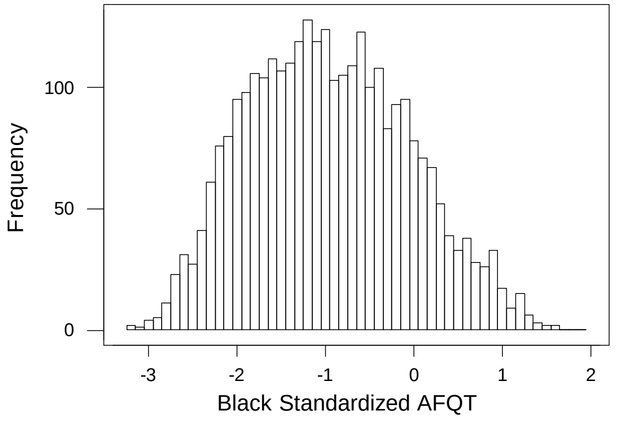

FIG. 4 Histogram of black AFQT scores standardized by age

which is decreasing in λ. For movers the between period average change is

$$E\left[\lambda p_2 + (1-\lambda)\, q\,(s, a(s)) - \lambda p_1 - (1-\lambda)\, q\,(s, a(s))\,|P_1 < \lambda p_2 + (1-\lambda)\, q\,(s, a(s))\right]$$

or equivalently

$$E\left[\lambda\left(\varepsilon_2 + \nu_2\right) - \lambda\left(\varepsilon_1 + \nu_1\right)|\lambda\left(\varepsilon_2 + \nu_2\right) - \varepsilon_1 > 0\right] = \frac{2\lambda\sigma_\varepsilon}{\sqrt{1+\lambda}}\left(\frac{2}{\sqrt{2\pi}}\right) \tag{49}$$

Expected wage gains to black movers are smaller than to white movers, but expected wage gains to black stayers are larger than to white stayers. A more subtle implication of equations (48) and (49) is that overall whites have larger average gains. To see this note that in equilibrium the probability that any individual moves is $\frac{1}{2}$. This is because firms correctly infer quality in equilibrium giving $pr(\text{move}) = \Phi\left(\beta\right) = \Phi\left(0\right) = \frac{1}{2}$. From this I can compute overall average wage in the second period.

$$E\left(w_2\right) = E\left[\lambda p_1 + (1-\lambda)\, q\,(s, a(s))\right] + \frac{1}{2}\frac{(1-\lambda)\,\sigma_\varepsilon}{\sqrt{1+\lambda}}\left(\frac{2}{\sqrt{2\pi}}\right) + \frac{1}{2}\frac{2\lambda\sigma_\varepsilon}{\sqrt{1+\lambda}}\left(\frac{2}{\sqrt{2\pi}}\right) \tag{50}$$

which reduces to

$$E\left(w_2\right) = Q\left(s, a\right) + \sqrt{\frac{\sigma_\varepsilon^2\left(1+\lambda\right)}{2\pi}} \tag{51}$$

Average second period wages increase with λ. Thus, controlling for education and ability, whites will earn more in the second period. If the relative gain to whites is large enough, the model even allows for average wages without controlling for education or ability to be larger for whites in the second period.

3. DATA

The data used for the empirical analysis come from two sources, the National Longitudinal Survey of Youth 1979 (NLSY79), and the Baccalaureate and Beyond Longitudinal Study of 1993 (B&B). The NLSY79 conducted yearly surveys of 12,686 individuals between 1979 and 1994. After 1994 surveys were conducted every other year. The individuals in the study were born between 1957 and 1964, making them between 14 and 21 years old in 1979. I only use the 1979 through 1993 surveys. The NLSY79 oversamples minorities, poor whites, and the military. As well as demographic variables, the NLSY79 collects detailed work history information and aptitude measures.

The NLSY79 sample used for the primary analysis is similar to that used by Farber and Gibbons (1996) and Oettinger (1996). The careers of military personnel may follow a different path from those of the general population, so I drop the military sub-sample. Women are more likely to have breaks in their work experience, so I limit my sample to men. The model makes predictions about wages at labor force entry and as experience is accumulated, but the data does not clearly show when a person moves into the labor force. I define the transition to the full-time labor force as in the aforementioned papers. First, an individual is said to have strong labor force attachment in any year during which they work at least half of the weeks and average at least thirty hours of work during those weeks. An individual is said to enter the labor force in the first year during which he has strong labor force attachment in a run of at least three consecutive years with strong labor force attachment. Individuals having strong labor force attachment in 1979 are eliminated because it is impossible to tell when they transitioned to the full-time labor force. Several individuals have multiple runs of strong labor force attachment of at least three years, broken up by one or more years without strong labor force attachment. For example a man might be employed for four years, then unemployed for one year, and then employed again for three more years. I use only the first run of employment for these individuals.

In 1980 11,914 of the NLSY79 respondents were given the Armed Services Vocational Aptitude Battery (ASVAB). This test consists of ten sections each measuring different skills. The Armed Forces Qualification Test is constructed from a subset of these. There are two different versions of the AFQT. The one used in this paper was defined in 1980. The raw AFQT score is created using the word knowledge, paragraph comprehension, arithmetic reasoning, and numerical operators section scores.[26] AFQT scores increase with age. To account for this I took the raw scores and standardized them for each age group. I did this using the entire sample of $11,878$ individuals who took the test under normal conditions, using the appropriate weights. Thus an individual who took the test at age fifteen and scored one standard deviation above the average raw score for a fifteen year old would be comparable to an individual who took the test at age twenty-three and was one standard deviation above the average of individuals who took the test at age twenty-three, even though the raw score of the fifteen year old would be lower.[27]

The NLSY79 records detailed information on up to five jobs for each individual. For job characteristic variables such as occupation and indus-

[26] The exact formula for this score is $AFQT80 = WK + PC + AR + \frac{1}{2}NO$.

[27] Another formula was used starting in 1989 to compute AFQT. This version uses the math knowledge section instead of the numerical operator section and has a slightly different weighting. These scores were also computed. The use of this alternative definition had little impact on the empirical findings.

try, I used the information on the job the individual reported as currently working in. For individuals who reported that they were currently working at multiple jobs, I used the first reported job. For individuals who did not report that they were currently working on any job, I took the information on the job they reported as their Current Population Survey (CPS) job. Hourly rate of pay is recorded for each job and this is used as the dependent variable (in logs) in my wage regressions. I adjusted the hourly rate of pay into real terms using the Consumer Price Index[28] reported by the Bureau of Labor Statistics.

For an individual's education I took the reported highest grade completed if it was available. For observations missing this variable I used the highest grade completed as reported in the previous year. For race I used race as reported by the interviewer unless it was not reported, in which case I used the sample to identify race. Because the model is based on the idea that employers take race as information when creating a pay scale this seems to be appropriate, although the self reported race might make more sense in the education decision. Using self reported race produces similar results. I restrict my sample to whites and blacks.

Finally, I drop observations missing key variables such as highest grade completed or hourly rate of pay. This leaves a sample of 2,985 individuals and 20,840 observations. Of the remaining individuals, 828 are black and 2,157 are white. The dramatic reduction in the number of individuals from the full sample (N=12,686) to the sample I use (N=2,985) comes primarily from 3 sources: the elimination of women, the elimination of the military subsample, and the requirement that men in the sample report at least three consecutive years of strong labor force attachment.

[28] The exact series ID is CUSR0000SA0

The B&B is a large nationally representative sample of college students graduating from college with a bachelor's degree in the 1992/1993 school year. The original sample was drawn from the National Postsecondary Student Aid Study (NPSAS) of 1993, collected by the U.S. Department of Education's National Center for Education Statistics (NCES). The initial sample contains 11,192 graduates from 649 universities. The baseline data comes from the NPSAS of 1993. Follow-ups took place between June and December of 1994, 1997, and 2003, approximately 1 year, 4 years, and 10 years out of college.

In the empirical work that follows I attempt to analyze the NLSY79 and B&B in as similar a manner as possible. Using the B&B data it is not difficult to determine when individuals transition to the labor force. I assume that all of the men in the sample working more than thirty hours a week in any sample year have entered the labor force. Once again I eliminate women from the sample.

The most notable difference between the NLSY79 and the B&B samples is the B&B's omission of a good control for ability. As indicated by the model in Section 2, controlling for ability is crucial. Luckily, the B&B does contain student transcripts including entrance exam scores. There is no single entrance exam score recorded for all individuals, but about 60% of the students have a recorded SAT score and 35% have an ACT score recorded. To create a single measure, I compute each student's percentile ranking on each exam they have recorded on their transcripts.[29] This gives me 8,954 individuals, or about 80% of the sample. College entrance exam scores do not provide as good a measure of ability as does AFQT. While the

[29] I created these percentiles based on my own sample, but also compared these to the percentile rankings for both the SAT and ACT reported by The College Board for 1989. The transformations were almost identical. In what follows I use the created transformation.

AFQT has been shown to be a racially unbiased measure of skill (Wigdor and Green, 1991), the SAT and ACT are likely not. As is discussed in Chapter 3, entrance exam scores may systematically underestimate the ability of black students. Also, the purpose of these exams is to measure aptitude rather than exogenous pre-labor market skill. In the empirical analysis I will use entrance exam scores as a measure of ability as there is not a better measure, but it is important to be mindful of their deficiencies.

The structure of the B&B removes the need for inclusion of an individual's highest grade completed, as all of the men in the sample have a college degree. I do include an individual's G.P.A. and a measure of the school quality to control for other educational aspects. The measure of school quality used is the average entrance exam score[30] at the school. This variable is empirically constructed for each school using the entire B&B sample and excluding the individual of observation.

The B&B reports detailed information on each individual's primary job in April of each survey year. One variable recorded is the job's start date. From this I am able to construct an accurate measure of tenure for each individual. I control for experience in a slightly different manner than in the NLSY79 sample. All of the students graduated in the 1992/1993 academic year, and most of them graduated in May of 1993. Because the follow-up surveys refer to the April job of the given survey year, all of the graduates have virtually identical experience in each survey year. For this reason I include indicator variables for year.[31] These year dummy variables take the place of the experience variables used in the NLSY79 sample. They allow for the estimation of a more flexible return to experience as well as

[30] As with the individual entrance exam scores, these are measured in terms of the entrance exam percentile.

[31] There are only three surveys in the B&B panel, 1994, 1997, and 2003. I include indicator variables for 1997 and 2003.

for the returns to variables such as G.P.A. and school quality to vary over time.

4. EMPIRICAL FINDINGS

The focus of this paper is discrimination in the labor market. I am primarily concerned with wage differences in a correctly specified model. Education is only of interest because of the implications on wages. For this reason the majority of the empirical work that follows will focus on wages. I will only briefly address education in order to justify my initial claim that blacks acquire more education than whites of a given ability level.

4.1. Education

Controlling for ability, blacks receive about one more year of education than do whites. Table 5 gives the results of a regression of an individual's highest grade completed on their race, AFQT, and other controls using the NLSY79. The amount of extra education received by black students varies by AFQT as is depicted in Figure 5. This figure gives the predicted education level by race and AFQT under the restricted specification as given in Table 5.[32] Recall AFQT has a standard normal distribution. Thus, a black individual is predicted to obtain approximately one more year of education than a white individual at the average level of ability. At approximately 1.3 standard deviations above the average ability level, the predicted education is the same for black and white students. This result is robust to model specification and consistent with model proposed in section 2. Recall that

[32] Using either of the other specifications produces similar results, the primary difference being that under the other specification, black and white education levels are also expected to be equal approximately $2\frac{1}{4}$ standard deviations below the average ability level.

at the highest and lowest ability levels ($\underline{a}$ and $\overline{a}$) there is no incentive for individuals to signal so that at these levels of ability black and white students invest in the same level of education.[33] It is also notable that the marginal effect of AFQT is positive for all reasonable values. This is consistent with the assumption that $s'(a) > 0$.

The sample for each regression in Table 5 consists of all black and white males from the NLSY79 that were not in the military subsample. Only individuals with valid responses for the reported variables were included, leading to the decreasing number of individuals as more variables are included. The dependent variable is highest grade completed. Each round of the NLSY79 records highest grade completed, and I used the largest of these values reported by each individual.

4.2. Wages

Regarding wages, the theory developed predicts that blacks will face a pay scale that places more weight on a firm's inference about quality than will whites. Thus, the return to education will be higher for blacks. Similarly, whites will have a higher return to ability.[34] With respect to changes over time the theory predicts that the measured return to experience will be higher for whites while the measured return to tenure will be higher for blacks. I examine these predictions using data from both the NLST79 and the B&B. Table 6 reports the results from the NLSY79 and Table 7

[33] The estimated model actually shows that at high levels of ability white students will get more education than black students. This may reflect the inability of the quadratic specification of AFQT to capture both its increasing marginal effect at low levels and the fact that at the highest levels of ability the marginal returns will approach zero (due to the maximum education level). Because there are many individuals with ability levels in the range of increasing marginal returns to ability, and few at the highest levels of ability, the estimate will primarily reflect the increasing marginal returns to ability.

[34] Chapter 3 presents results providing evidence that black college graduates face a higher return to school quality and a lower return to ability than do white college graduates.

TABLE 5
Regression of an individual's highest completed grade level on given variables.

	Specification		
	Restricted	School	School & Family
Black	1.02^{***} (0.09)	1.13^{***} (0.142)	1.47^{***} (0.183)
AFQT	1.91^{***} (0.035)	1.79^{***} (0.048)	1.57^{***} (0.064)
Black*AFQT	-0.539^{***} (0.114)	-0.687^{***} (0.162)	-0.511^{**} (0.117)
$AFQT^2$	0.302^{***} (0.027)	0.476^{***} (0.0396)	0.453^{***} (0.051)
Black*$AFQT^2$	-0.235^{***} (0.043)	-0.508^{***} (0.086)	-0.518^{***} (0.117)
Books in library (×1000)	–	0.008^{**} (0.004)	0.005 (0.005)
% School disadvantaged	–	−0.001 (0.002)	0.003 (0.003)
Average daily attendance (%)	–	−0.001 (0.003)	−0.001 (0.004)
% 10^{th} grade dropout	–	-0.008^{***} (0.002)	-0.006^{**} (0.002)
% Black students	–	-0.01^{***} (0.003)	-0.008^{*} (0.004)
% Black faculty	–	0.012^{**} (0.005)	0.004 (0.006)
% Teachers with adv. deg.	–	0.006^{***} (0.002)	0.004^{**} (0.002)
Teacher Starting Salary (×1000)	–	0.0 (0.0)	0.0 (0.0)
Mother's H.G.C.	–	–	0.034^{*} (0.021)
Father's H.G.C.	–	–	0.05^{***} (0.016)
Sibling's H.G.C.	–	–	0.131^{***} (0.021)
Constant	12.7^{***} (0.043)	12.3^{***} (0.458)	9.66^{***} (0.637)
N	4,939	2,375	1,497

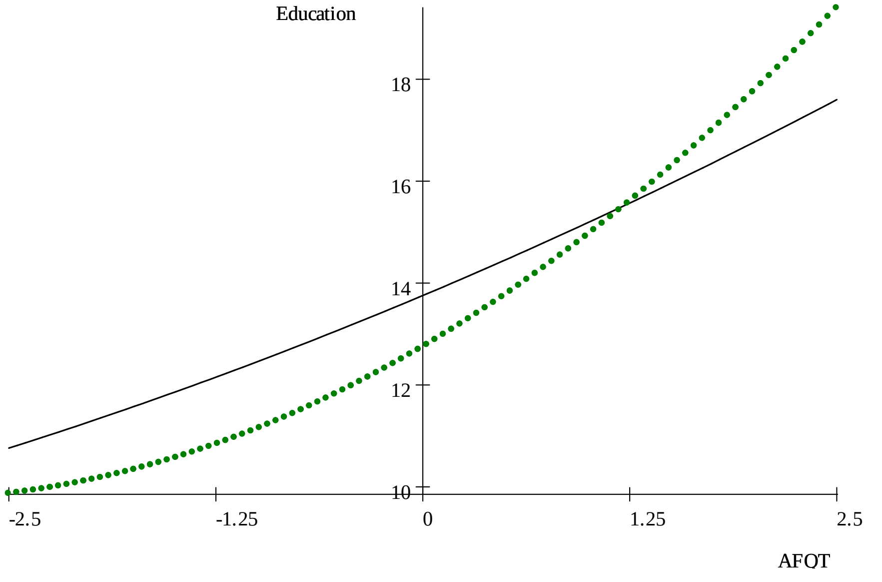

FIG. 5 ··· White predicted education; — Black predicted education

reports those using the B&B93. The dependent variable is the natural log of hourly wage.[35] Using each data set I separately examine both the entire sample and a subset of only the men in their early careers, those with five or fewer years of experience. I report both fixed effects and random effects estimates for each specification.

The direct effects of ability and race on wages can only be seen in the random effects model. Random effects estimation is only consistent if the unobserved heterogeneity is uncorrelated with included variables. This is a restrictive assumption but the inclusion of AFQT and college entrance exam score to control for ability removes much of what is typically considered unobserved and correlated with included variables. Also, the

[35]Explanatory variables included in the models but not reported are a set of 9 occupation indicator variables, 8 industry indicators, the age at which strong labor force attachment was first formed, and indicator variables for wether the individual lived in an urban area, is married, and is covered by collective bargaining. There are also a full set of interactions with the black indicator variable.

TABLE 6
Panel regression of the natural log of hourly wages on the given variables as well as the age at which an individual entered the labor force, indicator variables for marital status, whether the individual was covered by collective bargaining, whether the individual lived in an urban area, and a set of indicators for industry and occupation as well as their interaction with the black indicator.

	Full Sample		Experience≤5	
Panel Specification	FE	RE	FE	RE
Black	–	0.232**	–	0.1052
		(0.099)		(0.1068)
Education	0.0574***	0.0506***	0.038***	0.044***
	(0.0082)	(0.0033)	(0.0109)	(0.0035)
Experience	0.079***	0.0732***	0.0872***	0.0747***
	(0.0037)	(0.0036)	(0.0113)	(0.0112)
Tenure	0.0229***	0.0268***	0.0075	0.015***
	(0.0034)	(0.0033)	(0.0065)	(0.0056)
Experience^2	−0.0028***	−0.0026***	−0.0031*	−0.0027
	(0.0003)	(0.0003)	(0.0018)	(0.0018)
Tenure^2	−0.0018***	−0.0018***	−0.0011	−0.0004
	(0.0003)	(0.003)	(0.0011)	(0.0008)
Black×Education	−0.014	−0.0106	0.0227	−0.0065
	(0.0195)	(0.0066)	(0.0248)	(0.0069)
Black×Experience	−0.0241***	−0.0210***	−0.026	−0.0201
	(0.0072)	(0.007)	(0.0216)	(0.0215)
Black×Tenure	0.008	0.0066	0.0379***	0.0342***
	(0.0071)	(0.0067)	(0.0125)	(0.0113)
Black×Experience^2	0.0003	0.0003	0.0006	0.0005
	(0.0006)	(0.0006)	(0.0036)	(0.0036)
Black×Tenure^2	−0.0003	−0.0004	−0.0065***	−0.0064***
	(0.0007)	(0.0007)	(0.002)	(0.0018)
AFQT	–	0.0681***	–	0.0667***
		(0.0078)		(0.0081)
Black×AFQT	–	0.0147	–	0.0132
		(0.0147)		(0.0151)

TABLE 7
Panel regression of the natural log of hourly wages on the given variables as well as the age at which the individual graduated, a set of four indicators for region, a set of fifteen indicators for industry, a set of twelve indicators for occupation, and a set of six indicators of college major.

Years in panel	1994, 1997, 2003		1994, 1997	
Panel Specification	FE	RE	FE	RE
1997	0.3137*** (0.0655)	0.283*** (0.0627)	0.2736*** (0.0629)	0.2499*** (0.0577)
2003	0.8110*** (0.06645)	0.8231*** (0.0624)	–	–
Black	–	0.0330 (0.0555)	–	0.0206 (0.057)
Black×1997	−0.0435 (0.0506)	−0.0588 (0.0471)	−0.0855 (0.054)	−0.0856** (0.0450)
Black×2003	−0.0242 (0.0626)	−0.0583 (0.0542)	–	–
Tenure	0.004* (0.0025)	0.0079*** (0.0020)	0.0025 (0.0048)	0.0166*** (0.0033)
Tenure^2	−0.0001* (0.0001)	−0.001** (0.0000)	−0.0001 (0.0001)	−0.0002 (0.0001)
Black×Tenure	0.0066 (0.0158)	0.0107 (0.0136)	0.0271 (0.0326)	0.044** (0.0225)
Black×Tenure^2	−0.0011 (0.0008)	−0.0008 (0.0007)	−0.0022 (0.0018)	−0.003** (0.0013)
Entrance Exam	–	−0.003 (0.0003)	–	−0.0006** (0.0003)
Black×Entrance Exam	–	0.0011 (0.001)	–	0.0010 (0.0010)
G.P.A.	–	0.0730*** (0.0171)	–	0.0793*** (0.0163)
G.P.A.×1997	0.0039 (0.0196)	0.0105 (0.0189)	0.0183 (0.0187)	0.0179 (0.0173)
G.P.A.×2003	−0.0063 (0.0196)	−0.016 (0.0186)	–	–
School Quality	–	0.0021*** (0.0006)	–	0.0026*** (0.0005)
School Quality.×1997	0.0003 (0.0006)	0.0003 (0.0006)	0.0004 (0.0006)	0.0004 (0.0006)
School Quality×2003	0.0011* (0.0006)	0.0007 (0.0006)	–	–

estimated coefficients for the fixed effects and random effects models are similar on all of the variables except for the interaction of the race indicator and education in Table 6 and tenure in Table 7. Except for the coefficient on these terms, all of the coefficients from the random effect estimation are within one standard error of their corresponding fixed effects estimate. I will proceed referring to the random effects results from each table unless otherwise stated.

The theory predicts that upon labor force entry there will be no wage differential. This fact is supported by the B&B results from Table 7. In 1994, when all of the individuals have one year of experience, black graduates are expected to earn between approximately 2% and 3% more than white graduates with similar entrance exam scores, but the difference is not significant. The NLSY79 results do not fully support this prediction. When looking at only the early careers of individuals, there is no significant difference between initial wages of whites and blacks. When this analysis is extended to the full sample, the NLSY79 results indicate that *black* individuals actually earn significantly more than white individuals. One implication of the theory is that if ability is controlled for but education is not, then blacks will appear to earn more than whites. This stems from the fact that for a given level of ability black individuals will acquire more education than white individuals. The results using the NLSY79 full sample may arise because the included measure of education does not pick up all of the relevant aspects of education such as school quality and grades.[36]

As is implicitly assumed in the model, the returns to education, as measured but the highest grade completed for the NLSY79 and school

[36] An alternative explanations is that if a black and a white individual have the same level of education and AFQT, it is possible that the black individual has more positive unobservables.

quality as well as G.P.A. for the B&B, are positive. The model predicts that blacks should have higher returns to education than whites. This prediction is not supported by the data. According to Table 6 the estimated returns to education are lower for blacks than they are for whites, although the difference is not statistically significant. The model similarly predicts that blacks will have a lower return to ability. This prediction is also not supported by the data. Both the NLSY79 and the B&B estimate blacks as having higher returns to ability, although the difference is once again not statistically significant.

An important implication of the theory developed in Section 2 is that blacks will face a lower measured return to experience than will whites. This aspect of the model is generally supported by the data. When looking at the full sample of the NLSY79, the marginal return to experience for blacks is $0.0522 - 0.0046(experience)$ and the marginal return to experience for whites is $0.0732 - 0.0052(experience)$. The difference is statistically significant at all reasonable levels of experience. Figure 6 depicts the return to experience separately for backs and whites. The point estimates from the sample containing early careers are almost identical although the standard errors are larger, likely due to fewer observations of each individual.

With regard to experience, the B&B tells much the same story as the NLSY79. In 1997 all of the individuals in the sample have 4 years of experience, while in 2003 they have 10 years of experience. The estimates using the full sample indicate that in 1997 white individuals had wages that were approximately 28% higher than their 1994 level. By 2003 white wages were approximately 82% higher than 1994 wages. Relative to 1994, black wages had only risen by 22% in 1997 and 76% in 2003.[37] Although the point

[37] Black wages were approximately 3% higher than white wages in 1994, indicating black wages were not 6% lower than white wages in 1997 as the difference in return to

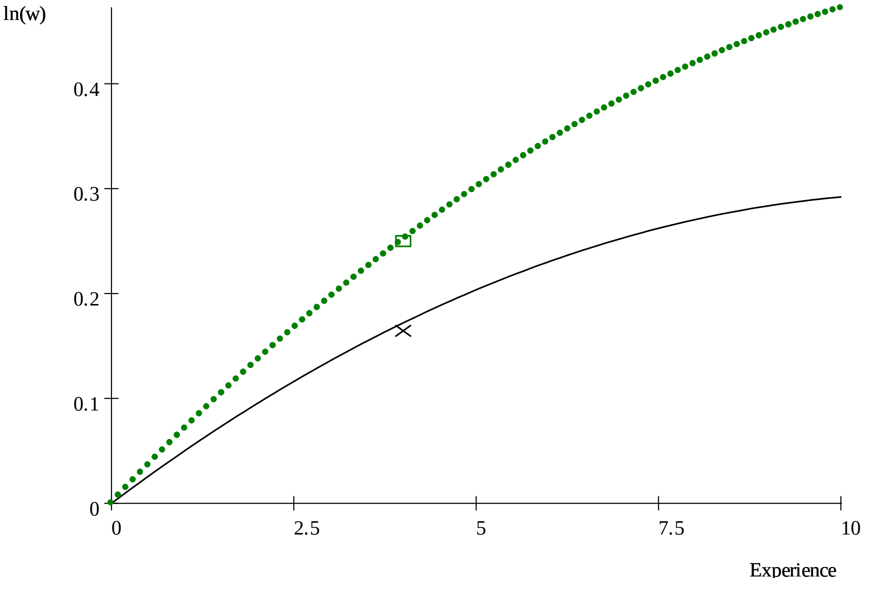

FIG. 6 $\cdots$ White return to experience; $-$ Black return to experience

estimates from the full sample indicate that blacks face a lower return to experience than whites, the difference in wages is not statistically significant in any year. Turning to the early careers in the B&B, the estimates indicate that the 1997 wages of white individuals are 25% higher than their 1994 wages, while black wages only rise by 16%. In these estimates the relative difference between white and black wages in 1994 and 1997 is statistically significant.[38] These estimates for the return to experience are similar in magnitude to those using the NLSY79. The square on Figure 6 represents the predicted wage growth for whites in due to increasing experience, while the "X" represents that of blacks.[39]

---

experience might indicate, but rather only 3% lower. Because we do not know wages in 1993 when experience was equal to zero it is hard to say what the 1^{st} year return to experience is.

[38] Although the relative difference between wages in 1994 and 1997 is statistically significant, it should be noted that white and black wages are not statistically different from each other in either year.

[39] These are not technically correct as we can not estimate the returns to the first year

TABLE 8
Estimated marginal effects of tenure on the natural log of wages.
Reported estimates are for early career sub-samples

	NLSY97	B&B
White	$0.015 - 0.008\,(tenure)$	$0.0166 - 0.0004\,(tenure)$
Black	$0.0492 - 0.0208\,(tenure)$	$0.0606 - 0.0064\,(tenure)$

While the model predicts that blacks will receive lower returns to experience, it also predicts that they will face higher measured returns to tenure. Both data sets provide some support for this prediction, although only when examining individual's early careers. Table 8 reports the marginal effects of tenure on the natural log of wages for men in their early careers. For both samples the coefficients on interaction of the black indicator and tenure as well as its quadratic are jointly significant.[40] Because tenure in this sample must be fewer than five years, the estimated marginal return to tenure for blacks is always greater than that of whites in the B&B sample. For the NLSY79 this is only true up to approximately 2.7 years of tenure. This may reflect the inability of the quadratic term to capture the flattening out of the return to tenure in the NLSY79 sample. The average tenure for this group is 1.9, with a standard deviation of 1.5. Thus, the point at which estimated marginal return to tenure for whites surpass that of blacks is greater than one standard deviation above the mean level of tenure. When looking at the full samples rather than the early career samples, the returns to tenure for blacks and whites do not differ in a statistically significant manner, although the point estimates do indicate that blacks receive a higher return to tenure.

of experience in this model. These returns actually represent the return to the second, third, and fourth year of experience.

[40] For the NLSY79 sample the coefficients have a p-value of 0.0013 on joint significance level. The p-value on the joint significances level in the B&B sample is 0.0589.

TABLE 9
Coefficient on black indicator variable for select model specifications

Included Explanatory Variables	Sample			
	experience=1	experience=6	year=1990 w/ job history	year=1990 w/o job history
education	−.058	−.099	−.074	−.066
	(.017)	(.033)	(.030)	(.029)
AFQT	−.009	−.016	.013	.025
	(.019)	(.032)	(.031)	(.032)
both	−.012	−.039	.011	.018
	(.019)	(.032)	(.031)	(.032)

Table 9 reports the coefficient on a black indicator variable under different model specifications using the NLSY79, and provides a clear way to examine which model specifications predict a wage gap. The row indicates whether education and ability explanatory variables are included.[41] The first two columns report the results of a wage regression using only observations of a given experience. The third and fourth column treat the data as cross-sectional observations taken in 1990, with column three controlling for experience and education. Column four represents a regression similar to that in Lang and Manove (2004) but with different results.

According to the theory, if ability is ignored blacks will appear to earn less, while if education is ignored whites will appear to earn less. The first prediction is supported by the data. The coefficient is statistically significant if and only if AFQT is not included. The second prediction is not supported. When education is not included the coefficient is not statistically significant, although including only AFQT does report blacks as doing

[41] Explanatory variables included in the models but not reported are a set of 9 occupation indicator variables, 8 industry indicators, the age at which strong labor force attachment was first formed, and indicator variables for wether the individual lived in an urban area, is married, and is covered by collective bargaining. Column 3 also includes experience, tenure and the squares of each. No interactions are included.

better relative to whites than does including education and AFQT.[42]

The theory also predicts that conditional on ability and education there will be no wage gap at entry into the labor force, but one will emerge over time. The prediction seems to hold when AFQT is excluded but not when it is included. Although the results are not reported for the regressions using sample with experience between one and six, the coefficients consistently decrease and remain statistically significant if AFQT is excluded. When AFQT is included the coefficient is never significant and fluctuates around zero. This contrasts the results of the random and fixed effects regressions reported in tables 6 and 7 which show a wage gap emerge as experience increases. The results in table 9 may miss this because of the restrictive form of the regressions.

5. CONCLUSION

The primary contribution of this chapter is to identify and model a single mechanism that explains both the relatively high investment in education by blacks and the black-white wage differential. It also highlights the importance of accounting for ability, education, and temporal factors when examining wages. The empirical results fit well with the model. I find that if ability and education are controlled for there is no wage gap at labor market entry under most specifications.[43] Because blacks face lower returns to experience, a gap emerges over time. Wages for whites grow by about 2% more than wages of comparable blacks each year, although this difference shrinks over time. This result is very similar to the results in

[42] When the sample is conditioned on experience=2,3,4, or 5 this result still holds. The coefficient is not significant but including only AFQT does report blacks as doing better than including both education and AFQT.

[43] Using the full sample of the NLSY79 I actually find that *blacks* earn more at labor force entry.

Oettinger (1996). Lang and Manove (2004) find using cross-sectional data and not accounting for experience or tenure that there is a black white wage differential. The model proposed here shows that the differential is due to lower rewards to experience for blacks as a result of the difficulty faced in determining good job matches.

As mentioned previously, one possible objection to the model is that it predicts that blacks will have higher initial wages than whites if ability has the same distribution for both populations. In my NLSY79 sample the average entry wage for blacks is about 1.2% lower for blacks than it is for whites. This can be explained by the fact that the average AFQT for blacks is about one standard deviation lower than that of whites. AFQT is likely influenced by pre-labor market factors that differ systematically by race. This does not affect the validity of the model as long as AFQT is exogenous to the individual. Because there is a difference in the distributions of AFQT scores by race, and because AFQT scores play a significant rose in determining wages, future research should be done into the specific factors causing this difference.

Chapter III

Labor Market Effects of Preferential College Admission

1. INTRODUCTION

Preferential college admission, commonly referred to as affirmative action, has a long history and an uncertain future. Most Americans believe that the goals of affirmative action policies are noble, but much disagreement remains concerning the efficacy and equity of the policies. Of crucial importance to the effectiveness of affirmative action policies is their impact on labor-market outcomes. University admission policies are not made in a vacuum. Changes in these policies not only affect the school population, but also the beliefs held by potential employers about the school's population. This paper examines the effects of a school's level of affirmative action on the wages of the schools' graduates.

Much research has been done concerning the effects of preferential college admission on the composition of a student population. Card and Krueger (2004) and Long (2004) explore how the end of affirmative action policies in California and Texas have altered the makeup of students applying to college. Dickson (2004) provides evidence that affirmative action policies not only affect where students apply, but also which individuals apply to college. A large volume of literature also focuses on the effects of affirmative action polices on student educational outcomes. Datcher Loury and Garman (1993,1995), Kane (1998) and Bowen and Bok (1998) examine the college performance of minority students as a function of school selectivity. The results of the three studies are conflicting and inconclusive, but do show that college performance (e.g. GPA) depends on school selectivity and race. These questions are important to institutions when determining their optimal admission policy but do not address the performance of graduates once they leave college.

The purpose of this chapter is to measure the effects of a school's preferential admission on the labor-market outcomes of its graduates. Specifically, I look at wages as a function of individual as well as school characteristics including a measure of affirmative action. Past research on affirmative action either focuses on the effects of affirmative action in education on educational outcomes or the effects of affirmative action in the labor market on labor-market outcomes. Little has been done linking school polices on affirmative action to labor-market outcomes. Some researchers such as Datcher Loury and Garman (1993, 1995) examine wages as a function of school selectivity and race from which they are able to make limited inference about the possible effects affirmative action. They examine the relationship between individual ability and school selectivity noting that under affirmative action black students of a lower ability may be admitted to more selective schools. This paper differs in that I account for affirmative action directly while holding ability, GPA, and school quality constant. If school selectivity is valuable to employers because it carries information about unobservable ability, then affirmative action may impact wages beyond the direct effect of having lower ability black students in more selective schools.

I find that the level of a school's affirmative action does affect wages. Specifically, I find that the marginal effect of graduating from a school with the average level of affirmative action compared to a school of the same quality with no affirmative action is an approximate 13% reduction in wages for black students one year out of college. This wage gap disappears by the time the individuals have been in the labor market for four years. The wage reduction due to affirmative action offsets approximately 75% of the return to an equivalent increase in school quality. This relationship is consistent with the hypothesis that affirmative action decreases the signaling value of school quality for blacks. There is no significant relationship between affirmative action and white wages.

It is important to note what is not within the scope of this chapter. I am not looking at the full effects of affirmative action. Specifically, I am not looking at whether individuals are better off with or without affirmative action. Some students may be able to attend college or may move to a higher quality university because of an affirmative action policy. These students may be better off under affirmative action than they would have been without it. What I can say is that they receive a lower wage than had they gone to a school of the same quality with no affirmative action; however, such a school may not be in the choice set of these individuals. This research highlights the fact that affirmative action lowers the value of school quality as a signal of student ability to employers.

The paper proceeds as follows: Section two presents a theoretical basis for the effects of affirmative action on graduate wages and discusses their implications. Section three explores the data form the 1993 class of the Baccalaureate and Beyond Longitudinal Study, with section four discussing the measurement of affirmative action. Section five presents empirical evidence based on data. The final section discusses implications of the research and concludes.

2. THEORY

Employers never have perfect information about the productivity of a potential employee at their firm. This idea is the basis for the theory that follows. Without perfect information employers will use easily observable characteristics when making hiring decisions and wage offers. This will be especially important for newly graduated employees as they have less information available on which to base expected productivity. Two easily observable characteristics are school quality and race. Due to pre-market factors, it is likely that the underlying distributions of true productivity of college graduates differ systematically between races even without affirmative action. The presence of affirmative action may serve to exacerbate the differences. Not only will it likely lower the expected productivity (conditional on school quality) of black graduates, but it will likely increase the variance of the distribution of productivity. Thus, affirmative action will have two effects: black graduates from schools with high levels of affirmative action will be treated as though they were from a lower quality school, and the relative importance of school quality as a signal will decrease.

Formally, assume there is a continuum of individuals, each with ability level a_i that does not vary by firm or school and is unobservable to everyone. True productivity of individual i at firm j is represented by $P_{ij} = a_i + \varepsilon_{ij}$ where ε_{ij} represents the quality of the firm-individual match[44] and is unknown to both individuals and firms. Although firms do not know a_i, they know the distribution of a_i in the population to be $N\left[\mu_a, \sigma_a\right]$. Employers receive two signals as to the true value of a_i. First, they receive a direct signal $g_i = a_i + \nu_i$ where $\nu_i \sim N\left[\mu_\nu, \sigma_\nu\right]$ represents an individual

[44] In the analysis that follows, the inclusion of ε_{ij} is not critical to the results. It is included here because it adds realism to the model and because it allows for straightforward extension of the model to other applications.

specific error term in the measurement of ability. This direct signal can be thought of as grades, a test given upon application, or an interview impression. Employers also observe the individual's school s_i. The labor market is competitive with risk neutral workers and firms. Thus, employers make a wage offer equal to the individual's expected productivity.

Ability a_i is also unobservable to schools. They do observe a test score $T_i = a_i + u_i$ where $u_i \sim N[\mu_u, \sigma_u]$ represents measurement error in the exam. This test can be thought of as the SAT or ACT and is unobservable to employers. Without fully developing the value function for schools, I assume they follow the acceptance rule: accept individual i if $T_i > \underline{T}$ where $\underline{T}$ is set by the university and represents the lowest acceptable test score. It can be thought of as an admission bar below which students are not accepted. I assume that a_i, ν_i, u_i, and ε_{ij} are independent.

This can be thought of as a school system where all individuals apply[45] to a single university and those with the best test scores are accepted. This may be a strong assumption as in reality there are many universities and applicants apply at multiple schools. By assuming that all students attend the highest quality school they are admitted to, the model extends readily to a market with many schools. Although a change in admission policies may change the applicant pool for a university, as shown by Card and Krueger (2004) and Long (2004), it is likely that the actual pool of potential students does not change. As the admission bar is raised, fewer low score students may be observed applying, but this is because their probability of acceptance is low, not because they are out of the market for that school. A more troubling problem is the top students. If the admission bar is lowered the university becomes less selective, which may cause them

[45] This implicitly assumes that the value of college is greater than the cost for all individuals.

to switch to the university that was previously their second choice. The many factors other than selectivity entering into the school choice decision for students at the top of the school's ability distribution will likely make this group of students small.

Given the above parameters, employers will make a wage offer of $w_i = E[P_{ij}|s_i, g_i]$. Employers receive no signals as to the value of ε_{ij} and I assume $E[\varepsilon_{ij}] = 0$. Thus the wage offer is equivalent to

$$w_i = E\left[a_i | a_i + u_i > \underline{T}\ ,\ a_i + \nu_i = g_i\right] \tag{52}$$

For simplicity I will continue without subscripts. From equation (52) a, $a + u$, and $a + \nu$ have a trivariate normal distribution. Define a' as a normally distributed random variable with mean $\mu_a + \lambda(g - \mu_a - \mu_\nu)$ and variance $(1 - \lambda)\sigma_a^2$ where $\lambda = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\nu^2}$. It can be shown that a' represents the distribution of a conditional on $g = a + \nu$. Thus equation (52) is equivalent to

$$w = E\left[a' | a' + u > \underline{T}\right] \tag{53}$$

Without loss of generality assume that $\mu_\nu = 0$. Now a' and $a' + u$ have a bivariate normal distribution with a covariance of $\sigma_a^2(1 - \lambda)$. The wage offer becomes[46]

$$w = \lambda g + (1 - \lambda)\mu_a + \frac{\sigma_a^2(1 - \lambda)}{\sqrt{\sigma_a^2(1 - \lambda) + \sigma_u^2}} \left[\frac{\phi\left(\frac{\underline{T} - \lambda g - (1-\lambda)\mu_a - \mu_u}{\sqrt{\sigma_a^2(1-\lambda) + \sigma_u^2}}\right)}{1 - \Phi\left(\frac{\underline{T} - \lambda g - (1-\lambda)\mu_a - \mu_u}{\sqrt{\sigma_a^2(1-\lambda) + \sigma_u^2}}\right)} \right] \tag{54}$$

To simplify notation, define $x = \frac{\underline{T} - \lambda g - (1-\lambda)\mu_a - \mu_u}{\sqrt{\sigma_a^2(1-\lambda) + \sigma_u^2}}$ and $\Psi(x) = \frac{\phi(x)}{1 - \Phi(x)}$. The first two terms in equation (54) look like the standard wage offer

[46] See Greene (2003) p. 781.

from the statistical discrimination literature. They show that wages are a weighted sum of one's own productivity signal and the population average. The third term is more complex and represents the benefit of attending a more selective school conditional on grades and the population mean. It accounts for the skew and the shift in the distribution of ability created by the acceptance rule.

Now it is possible to examine the effects of affirmative action on wages. First note that $\lim_{x\to-\infty}\Psi(x)=0$ and $\frac{\partial\Psi(x)}{\partial x}>0$ for all x. For students graduating from a school that admits anyone, the third term in equation (54) disappears and their college diploma adds nothing to the wage. The higher a school sets its admission bar, the higher the wages of the school's graduates. If a school has two different admission bars, one for black students $\left(T^b\right)$ and one for white students (T^w) employers will be able to use race as another signal about true worker productivity. I will use differing admission bars as a theoretical definition of affirmative action. If $T^b<T^w$ then black students will receive lower wages than white students from the same school. What is even more disturbing is that this relationship holds even after grades have been controlled for. Thus a black graduate will earn less than a white student from the same school with the same GPA. This relationship arises from the fact that g is not a perfect predictor of a, causing employers to use additional information to estimate productivity. If employers were allowed to see individual entrance exam scores, affirmative action would not have an impact on wages. Instead, they can only observe the characteristics of the population of graduates. Therefore a lower admission bar hurts the students who would have graduated without the lower bar because it devalues their degree. In essence, employers "undo" the effects of affirmative action. They treat each university as two separate

universities, one for black graduates and one for white graduates.

It is important to consider the reasons for affirmative action. I assume throughout this model that a and ν are distributed identically for blacks and whites. This means that conditional on ability, black and white students have the same expected signal g of ability observed by employers, and that the distribution of true ability does not differ by race. If schools must implement affirmative action plans to maintain a representative population of minority students, it must be that the admission test scores differ by race. This difference enters the test scores through the distribution of measurement error for admission tests (u). I assume that $u^w \sim N\left[\mu_u^w, \sigma_u^w\right]$ and $u^b \sim N\left[\mu_u^b, \sigma_u^b\right]$ represent the distribution of test measurement error in the white and black populations respectively. Whether or not a school's affirmative action policy will be successful in correcting for the differences in u^w and u^b will depend on how exactly they differ.

If it is simply the case that $\mu_u^b < \mu_u^w$ while $\sigma_u^b = \sigma_u^w$ then a university can maintain affirmative action without any negative impact[47] on its graduates wages. By setting T^w and T^b such that $T^w - \mu_u^w = T^b - \mu_u^b$ the university can perfectly correct for the bias in the entrance exam. One can see from equation (54) that the admission bar and the mean error on the admission exam enter the wage equation in exactly the same manner, making the admission bar an effective tool for maintaining equal distributions of ability (a) between races.

Alternatively, if entrance exams are less precise for black students than they are for white students, the university will not be able to equate the distributions of ability (a) between races. This is the standard assumption in the statistical discrimination literature that $\sigma_u^b > \sigma_u^w$ and $\mu_u^b = \mu_u^w$. Be-

[47] Note that under this regime black graduates will actually have a higher wage than white graduates in the absence of affirmative action.

cause $\frac{\partial w}{\partial \sigma_u} < 0$ black college graduates will receive a lower wage than white college graduates even when there is no affirmative action $\left(T^w = T^b = \underline{T}\right)$. Lowering the admission bar for black students will only exacerbate this problem.[48] A lower admission bar and less precise entrance exam have the same effect on the distribution of true ability a at the university. They both decrease the average ability and increase the variance[49] of ability.

Having $\sigma_u^b > \sigma_u^w$ decreases wages in two ways. First, $\Psi(x)$ is decreasing with respect to σ_u^b. This represents the "bump" in wages received by graduates for attending a more selective school. At higher levels of σ_u^b applicants who are truly qualified in terms of ability are more likely to be rejected because of a negative realization of T_i, while less qualified applicants are more likely to be accepted because of a high realization of T_i. This serves to lower the expected value of true ability among those accepted. Second, σ_u^b makes school selectivity a worse indicator of true ability. This ensures that employers will put less weight[50] on $\Psi(x)$, further decreasing wages. This can be seen through the negative relationship between σ_u^b and $\frac{\sigma_a^2(1-\lambda)}{\sqrt{\sigma_a^2(1-\lambda)+\sigma_u^2}}$.

When $\mu_u^b < \mu_u^w$ there is a need at universities for affirmative action as black students will on average do worse on the entrance exam than will white students. Under these circumstances black students will be under represented at the university relative to the population of applicants. As

[48] This is only a "problem" when looking at the wage differential between black and white college graduates. Lowering T^b will lower black college graduate wages, but will also increase the number of black college graduates and the average quality of school that they attend. Efficiency questions about the policy are beyond the scope of this paper.

[49] It can be shown that

$$VAR\left[a'|a'+u>\underline{T}\right] = \sigma_a^2(1-\lambda)\left[1 - \frac{\sigma_a^2(1-\lambda)}{\sigma_a^2(1-\lambda)+\sigma_u^2}\Psi(x)\left(\Psi(x)-x\right)\right]$$

This variance term is decreasing in x, while x is increasing in $\underline{T}$ and decreasing in σ_u^2. Thus $VAR\left[a'|a'+u>T^b\right] > Var\left[a'|a'+u>T^w\right]$.

[50] In essecne, $\Psi(x)$ represents the skewness of ability caused by selectivity. Increasing σ_u decreases the skew and puts more weight on the alternative of 0.

mentioned previously, universities can perfectly correct for this if it is the only difference between races. The problem occurs when entrance exams are both lower on average and less precise for black students. If both $\mu_u^b < \mu_u^w$ and $\sigma_u^b > \sigma_u^w$, the university is able to correct for the difference in mean errors, but only at the cost of lowering the expected ability of black graduates, thus lowering their wages. Using affirmative action the university can admit enough black students such that the accepted population matches the applicant pool, but doing so will exacerbate the problems associated with less precise entrance exam scores.

The model provides several relevant empirical implications. First, black graduates will receive lower wages than white graduates from the same school with the same ability level and GPA if admission exams are less precise for blacks. This wage differential will be larger the more pronounced the school's affirmative action policy. If employers learn about true ability over time, then the wage penalty for going to a school with affirmative action should decrease over time. The model also implies that black graduates should face a lower return to school quality as affirmative action increases.

3. DATA

The data used come from the Baccalaureate and Beyond Longitudinal Study of 1993 (B&B). The B&B is a large, nationally representative sample of college students graduating from college with a bachelor's degree in the 1992/1993 school year. The original sample was drawn from the National Postsecondary Student Aid Study (NPSAS) of 1993, collected by the U.S. Department of Education's National Center for Education Statistics (NCES). All monetary variables are adjusted using the consumer price index[51] retrieved from the Bureau of Labor Statistics.

The B&B is unique in that it follows a large number of college graduates over an extended period. It contains detailed information on each individual's college experience, institution, and post-college work experience. The initial sample contains 11,192 graduates from 649 universities. The baseline data comes from the NPSAS of 1993. Follow-ups took place between June and December of 1994, 1997, and 2003, approximately 1 year, 4 years, and 10 years out of college.

What makes the B&B truly valuable is the level of detail with which the individual's college experience is recorded. Respondents were asked a broad range of questions about their university in 1993 and 1994. Even more important, each individual's transcripts were recorded. These give the graduate's GPA, major, and any admission test scores such as the SAT or ACT.

The B&B also contains a great deal of institution level data. It has a demographic breakdown, graduation rate, research classification, whether the university is public or private, and many other key institution level variables. These data combined with the transcript data allow for precise measurement of a student's educational achievement and outcome. Typically the researcher is only able to control for years of education, but with the B&B this is automatic as the entire sample has the same years of education,[52] at least initially. I am able control for much more of what an employer observes. Any subsequent degrees are also recorded.

Concerning labor force experience, the B&B contains information on the employment (full time, part time, unemployed, out of labor force) status of

[51] The CPI used 1982-1984 as the base period.

[52] There are actually a few individuals for whom the degree of record is their second bachelor's degree. I control for this in the empirical work that follows.

an individual for every month from graduation through the 1997 interview date. It also records the start date of any job held at the interview date. Together these allow me to create accurate tenure and post-college experience variables. Graduate enrollment is also recorded for every month, so I am able to control for subsequent education and determine whether an individual is primarily an employee or a student.

It is important that I be able to measure ability. There is no measure of ability recorded for all individuals, but about 60% of the students have a recorded SAT score and 35% have an ACT score recorded. These measures correspond well to T_i from the theoretical model of the previous section. To have a single measure of ability, I compute each student's percentile ranking on each exam they have recorded on their transcripts.[53] I then create a variable using this value for students that only took one of the exams. For students who took both exams (approximately 13%) I use the higher of the two percentiles, as this is likely what admission decisions are based on. For students with no test score reported on their transcripts, I use self reported test scores if they exist. Using this method I am able to assign a percentile score to 8,954 individuals, or about 80% of the sample.

4. MEASURING AFFIRMATIVE ACTION

The biggest hurdle to examining affirmative action is the difficulty faced in measuring affirmative action. Affirmative action is not a binary policy. Most schools have some form of a preferential acceptance policy, whether formal or not. The difficulty becomes quantifying the degree of preferential admission. For example, in 1996 the 5^{th} Circuit Court ruled that the use

[53] I created these percentiles based on my own sample, but later compared these to the percentile rankings for both the SAT and ACT reported by The College Board for 1989. The transformations were almost identical. In what follows I use the created transformation.

of race in admission was not permissible in the University of Texas system. With this ruling the University of Texas ended the explicit use of race in admission decisions, but one year later the Texas legislature implemented a plan guaranteeing admission to any public university in Texas to anyone graduating in the top 10% of their high school class. The goal of this plan was the same as that of the affirmative action plan in place before it. For analytic purposes the new plan should also be treated as affirmative action. The question then becomes: How preferential is a preferential admission plan?

As the above example shows, part of the difficulty comes from the ambiguous definition of affirmative action. Ideally I would like to consider any difference in the probability of admission of observationally equivalent individuals between races to be affirmative action. Summing the difference in probability of admission over all observations would then create an affirmative action score for each school. Figure 7 depicts this for a hypothetical school using only the SAT for admission decisions. Using this definition, affirmative action would be defined by the shaded area. Without a definition such as this, it is difficult to move beyond case studies as different schools' admissions policies are not comparable.

Unfortunately, the B&B does not have information on applicants who were not accepted. Thus a working definition of affirmative action must be developed. I use the difference in the average admission exam percentile between blacks and whites for all of the students at a school except for the individual of observation. Schools at which the average black score is above the average white score have their affirmative action measure set to zero as it is unlikely that these schools engage in reverse affirmative action. These schools represent approximately 10% of the schools in the

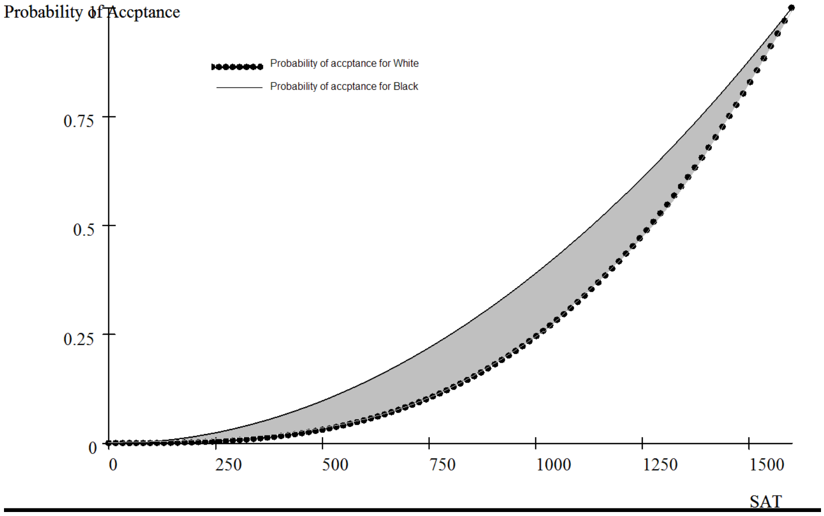

FIG. 7 Ideal measure of affirmative action. The shaded area represents the level of affirmative action at an institution.

survey. Empirically, the results in the following section are not sensitive to this censoring. This measure is more crude than the ideal measure, but it does provide information about the level of preferential admission. Figure 8 gives the distribution of affirmative action for individuals without censoring negative values to zero.

The B&B has several variables that control for, at least in part, school quality. One such variable is a control for the Carnegie classification of the institution as Research I, Research II, Masters/Bachelor's granting, or Liberal Arts. There is also a control for whether the university is public or private. To augment these, and to approximate the theory more closely, I also use a constructed measure of quality. The measure of quality I use is the average percentile on admissions exams for all students. Figure 9 gives the distribution of school quality. As with the affirmative action variable, this is constructed excluding the individual of observation. This is included

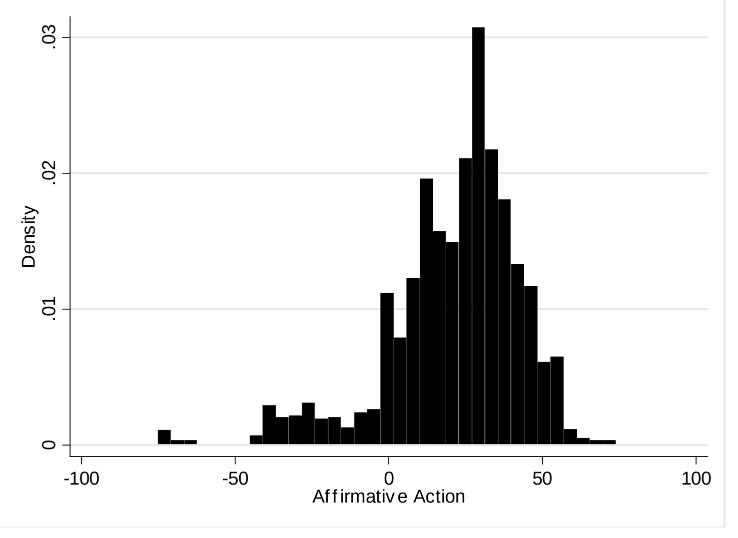

FIG. 8 The distribution of affirmative action. For the empirical work that follows, all negative values of affirmative action (approximately 10%) are set to zero.

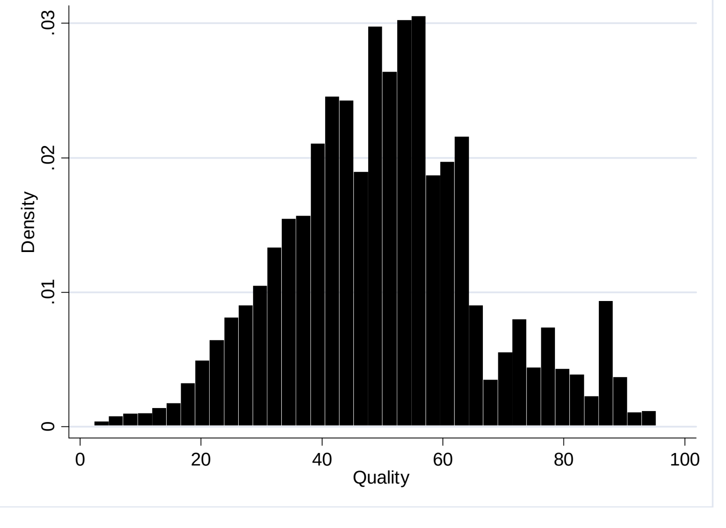

FIG. 9 The distribution of school qulity over individuals.

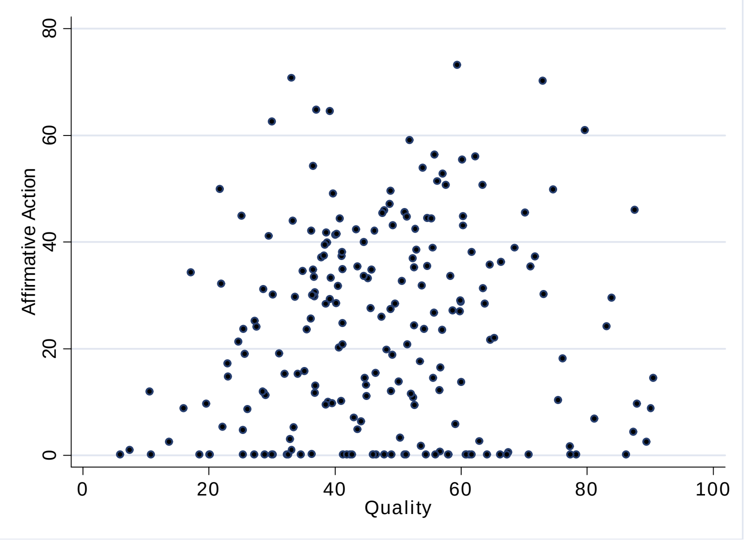

FIG. 10 Quality plotted against affirmative action

to represent what employers expect the average ability level to be for the graduates of a school. It also ensures that the created affirmative action variable does not appear significant simply because it is correlated with quality.

Figure 10 plots affirmative action against quality using these measures. From this figure it does not appear that quality and affirmative action are correlated. A regression of *aa* on *quality* using schools as the unit of observation gives a coefficient of 0.09 with a standard error of 0.07.

Tables 10 and 11 give the means and standard deviations of some variable of interest by observation and by school respectively. Note that in Table 10 the average school quality is similar for black and white students but the average entrance exam score is significantly lower for black students, as is GPA. The level of affirmative action is roughly the same for black and white students.

I am only able to construct a measure of affirmative action for 216 of the 649 schools in the B&B. To measure aa I must observe at least 3 observations from a school, one of which must be black and one of which must be white. Table 11 compares schools with at least one individual in the final sample to all of the schools in the B&B. It shows that the schools in the final sample tend to be slightly larger than the average school. For this reason I am slightly more likely to include public schools. Except for these variables, the schools included in my final sample resemble the schools of the full sample.

Note that the theory from section 2 refers to college admission, while in the data I only observe students graduating. This should not pose a problem as I am in essence looking at effective affirmative action. I am concerned with the individuals who have a diploma and what that diploma says about their ability. If a school admitted black students with a lower test score, but then failed out anyone below a certain ability level regardless of race, this school would not be considered to have an affirmative action program. The graduates of the school would all have the same expected ability. An affirmative action program that does not change the composition of a school's graduates should not be considered affirmative action in relation to the model presented in section two. Such a plan would not change the signaling value of a college degree. Studying the effectiveness of admissions policies at changing the makeup of graduates is interesting and useful to policy makers, but it has no bearing on the question addressed by this paper.

TABLE 10
Mean value of given variables by observation. Standard deviations are given in parenthesis.

		Sample		
		Full	White	Black
	1993 Hourly Wage	10.30 (6.39)	10.21 (5.74)	11.36 (11.57)
	1994 Hourly Wage	10.44 (5.86)	10.40 (5.48)	10.99 (9.21)
	1997 Hourly Wage	15.20 (10.15)	15.23 (9.98)	14.87 (12.10)
	2003 Hourly Wage	26.63 (20.83)	26.87 (21.31)	23.67 (13.07)
	GPA	2.98 (0.49)	3.0 (0.48)	2.67 (0.43)
	Entrance Exam Percentile	49.4 (27.8)	50.1 (27.3)	31.9 (27.2)
	Black	8.1% –	0 –	100% –
	Male	44.4% –	45% –	35% –
	Age Entered College	18.1 (1.5)	18.1 (1.5)	17.9 (1.5)
School Variables	% Black	7.8 (8.0)	7.4 (6.1)	13.1 (18.5)
	Affirmative Action	23.9 (15.2)	24.1 (15.3)	21.5 (14.8)
	Quality	50.1 (14.2)	50.2 (13.9)	49.4 (17.9)
	N	1813	1667	146

TABLE 11
Mean value of given variables by school. Schools in the included sample have at least one observation included in the empirical work that follows. The number of valid observations represents the number of schools with at least one individual observation of the given variable.

	Sample		# Valid
	Included	Full	In Full Sample
Individual Observations	25 (18.4)	17.2 (15.6)	649
White Observations	20.4 (16.2)	14.4 (13.9)	649
Black Observations	2.3 (2)	1.2 (2.2)	649
Avg. Percentile	47.5 (17.2)	46.7 (17.9)	621
Avg. White Percentile	51.2 (18.1)	49.5 (17.9)	585
Avg. Black Percentile	31.8 (25.7)	31.5 (25)	242
Avg. GPA	3.1 (0.2)	3.1 (0.3)	648
Avg White GPA	3.1 (0.2)	3.1 (0.3)	618
Avg. Black GPA	2.7 (.4)	2.8 (.4)	305
% Black	12.3 (17.7)	10.1 (17.7)	642
% White	74.6 (20)	75.6 (17.7)	642
Public	%60 –	%50 –	648
Enrolment	11,413 (10,280)	8,331 (9,920)	641
Affirmative Action	23.1 (18.7)	23.1 (18.7)	216
N	216	649	

5. EVIDENCE

To determine the effects of preferential admission on wages, I first regress the natural log of hourly wage on demographic and school level variables.

$$\ln(hourly\ wage_{it}) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2it} + e_{it} \qquad (55)$$

Equation (55) gives the empirical specification used, where X_1 represents the school level variables, X_2 represents individual level variables, and e is the error term. The equation is estimated separately for 1994, 1997, and 2003,[54] as denoted by the subscript t. The variables included in X_1 are aa, $quality$, the percentage of the student population that is black ($\%\_black$), an indicator for whether the school is public, and a set of indicators of the Carnegie classification of the school.[55] The percentage of the student population that is black, aa, and $quality$ are also interacted with an indicator $black$ assigned to black students.

The vector of individual variables X_2 contains $black$, an indicator for gender ($male$), GPA, entrance exam percentile ($percentile$), tenure and its square, post college work experience and its square, age and its square, and an indicator for whether the student was a resident of the school's state. Also included were sets of indicator variables to control for college major, any previous degrees, any subsequent degrees, and region.[56] Interactions

[54] Data were collected between June and December of each survey year. The 2003 survey asked about the individual's current job wile the 1994 and 1997 surveys asked about the primary job held in April of that year. Thus the years given in Table 12 refer to approximately 1 year, 4 years, and 10 years after graduation.

[55] The Carnegie classification used breaks schools into four categories, Research I, Research II, BA/MA Granting, and Liberal Arts.

[56] Four indicators were used to control for both previous and subsequent degrees, one each for another bachelor's degree, a masters degree, a Ph.D., and professional degrees. Six indicators were used to control for college major and five were used for region. One of each was of course omitted from all regressions.

of *black* with *percentile*, *GPA*, *tenure* and its square, and *experience* and its square were also included.

The B&B collected data in 1993 as well as 1994, 1997, and 2003, but I do not use the 1993 data. Instead, I use the 1994 data as the first labor-market observation. There are two primary reasons for this. First, because of the collection procedure, the length of time between graduation and the 1993 observation is more variable than that of 1994. Each surveys was collected between June and December of the given year. While the 1994 survey refers to the job held in April,[57] the 1993 data refer to the job held at the time of the interview. Thus, in the 1993 survey some individuals had been out of college for up to one year and four months, while others had only been out of school for one month. Because the 1994 survey refers to the April job, all individuals had been out of college for between one year and one year and 8 months. The second advantage provided by the 1994 observation is that by the time people have been out of college for one year, they have typically moved into the job that will become their career. In the empirical work that follows, I restrict the sample to individuals not working in a job they held during college. I do this because I am interested in the labor market for college graduates, and such individuals are likely not in that market.[58] Even after limiting the data in this manner, it is likely that in the 1993 data some individuals are working full time in new

[57] The 1997 survey also refers to the job held in April. The 2003 survey refers to the job held at the time of the interview just as the 1993 survey. I continue to use the 2003 data because it should not suffer from the problems associated with the 1993 data. As the graduates have been out of school for approximately ten years, it is unlikely that the difference in the length of time since graduation will have an effect on the results.

[58] There are two likely reasons we observe individuals not leaving a full time job they held during college. First, some students may continue working full time at jobs they used to pay for college while they search for jobs that require their degree. Second, there may be some students who have a developed career and go to college because it enhances their career. There will likely be little signaling value to a college degree when the graduates already have an established relationship with their employer.

jobs but still looking for a job in the field they prefer.[59] By April of 1994 all of the graduates have had at least one year to move into a job that will likely become their career. For these reasons I treat the 1994 observation as the initial labor market observation.

Table 12 gives the results of estimating equation (55) using ordinary least squares. Asterisks indicate significance with "*", "**", and "***" indicating that we can reject the hypothesis that the coefficient is zero at a 10%, 5%, and 1% significance level respectively. The sample for each year is restricted to those individuals working at least thirty-five hours each week. As mentioned in the preceding paragraph, I also restrict the sample to individuals not working at a job they held during college. I am able to do this because the B&B records the start date of all jobs on which it reports. This primarily affects the data from the 1993 survey, but a small subset of this group were still in their pre-degree jobs in April of 1994. Virtually all individuals changed jobs before April 1997.

Individual wages in 1994 provide the most insight into the questions addressed by this paper. By the time the individual has been out of college for 4 years, the market has provided a great deal of information about true ability and the value of the signal sent by school level variables has diminished. This is supported by Table 12, which suggests that college level variables become less important over time, while individual level characteristics become more important. The coefficients on the college level variables and their interaction with the race indicator decrease in both magnitude and significance over time, while the opposite trend is seen in individual level characteristics such as GPA[60] and *male*. In what follows I focus on

[59] An example of this would be a graduate with a degree in engineering who does not have a job offer upon graduation and so takes a job waiting tables while continuing to search for a job in engineering.

[60] The fact that college GPA becomes more important over time might indicate that

TABLE 12
Ordinary Least Squares regression results of the natural log of hourly wages for the given year on the X1 and X2. Standard errors are reported in parenthesis. Included variables that are not reported are: tenure and its square as well as their interaction with the black indicator, experience and its square as well as their interaction with the black indicator, age and its square, a set on indicator variables for previous degrees (5), a set of indicator variables of subsequent degrees (5), an in-state indicator, a set of indicator variables for region (5), a public school indicator, a set of indicators for the school's Carnegie Classification (4), and a set of indicators for college major (6).

		1994	1997	2003
University Variables	*aa*	0.0008	0.0007	0.0012*
		(0.0006)	(0.0005)	(0.0006)
	quality	0.0024***	0.002**	0.0012
		(0.0009)	(0.0009)	(0.001)
	% black	0.0046***	0.0035**	−0.0005
		(0.0015)	(0.0014)	(0.0017)
Individual Variables	black	0.1784	0.3019	0.398
		(0.2377)	(0.2839)	(0.5732)
	male	0.0412**	0.0801***	0.1825***
		(0.018)	(0.0173)	(0.0196)
	GPA	0.0249	0.0647***	0.0316
		(0.0202)	(0.0187)	(0.022)
	percentile	0.0003	0.0002	0.0005
		(0.004)	(0.0004)	(0.0004)
Black Interactions	*aa*	−0.0064***	−0.0019	−0.0001
		(0.0024)	(0.0019)	(0.0023)
	quality	0.0051***	−0.0026	−0.0015
		(0.0025)	(0.0023)	(0.0026)
	% black	−0.0035	−0.0041**	0.0007
		(0.0023)	(0.0021)	(0.0026)
	GPA	0.0061	−0.0175	0.0541
		(0.0734)	(0.0642)	(0.0743)
	percentile	−0.0045***	0.0007	0.001
		(0.0016)	(0.0013)	(0.0017)
	N	1817	2108	2284
	R^2	0.19	0.19	0.23

the 1994 results.

The variables of primary interest in Table 12 are *aa* and the interaction of *black* with *aa*. These measure the full effect of being in a school with a higher level of preferential admission on wages. For straightforward interpretation, Table 13 reports the marginal effects of school quality and affirmative action on the wages of white and black students separately. The effect of affirmative action on white wages is only marginally significant in 2003, and even in this year the magnitude is small. For black graduates, the effect is relatively large and significant in 1994 but disappears by 1997. To get a feeling for the size of the impact on wage, note that the average value of *aa* is 24. Thus, the marginal effect of affirmative action on black graduates of -0.0056 in 1994 indicates that one year out of college there is an approximate 13% wage differential between black individuals from schools with no affirmative action and black individuals from schools of the same quality with the average level of affirmative action.[61]

The coefficient on *quality* in 1994 is 0.0024, with black graduates realizing an extra return of 0.0051. The total return to school quality realized by blacks is 0.0074. Quality is measured as average entrance exam percentile at the university, giving it the same units of measurement as *aa*. For black students, the loss in wages due to a 1 point increase in affirmative action is approximately 75% of the gain in wages due to a one point increase in quality. This finding lends strong support to the model. It indicates that black wages depend primarily on the average ability level of the black pop-

grades not only signal ability but also of effort. It does not seem likely that the return to grades increases over time. It seems more likely that *GPA* is picking up people who work hard in college, and then continue to work hard in the labor force.

[61] The marginal effect of -0.0056 is robust to changes the variables included in X_1 and X_2. It remains significant and fluctuates between -0.005 and -0.007 depending on the other variable in the regression. It is even robust to the inclusion of the sets of indicators for industry, occupation, and major.

TABLE 13
Estimated marginal effects of affirmative action and school quality on the natural log of hourly wages for white and black graduates.

		1994	1997	2003
White Students	*aa*	0.0008	0.0007	0.0012*
		(0.0006)	(0.0005)	(0.0006)
	quality	0.0024***	0.002**	0.0012
		(0.0009)	(0.0009)	(0.001)
Black Students	*aa*	−0.0056**	−0.0012	0.0011
		(0.0023)	(0.0018)	(0.0022)
	quality	0.0074***	−0.0006	−0.0003
		(0.0024)	(0.0023)	(0.0025)

ulation at a given school, with the school's overall average ability playing only a minor role.[62] For example, say there are two schools. School A has no preferential admission and the average entrance exam percentile of 50 for both black and white students. School B admits white students such that their average entrance exam percentile is 69, but preferentially admits blacks such that the average for black students is 44 ($aa = 25$). According to these findings, the black students at both schools should receive the same wage offer conditional on their other characteristics. The black students at school B receive a small benefit from going to a school that is more selective for white students, but they are not treated by employers as having come from the same population as the white students from the same school.

It should be noted that the although affirmative action offsets most of the returns to school quality for black students one year out of college, the relationship does not hold at later dates. As expected, the wage penalty

[62] This relationship only strictly holds if the black population is small enough that it does not change the overall average entrance exam score for the school. This may not be true in the sample due to the construction of *aa*. If the black population is large enough that it has a significant impact on *quality*, then the loss due to affirmative action will actually be larger than 75% of the gain to quality. To see his note that if the average black entrance exam percentile drops, not only will *aa* increase, but *quality* will decrease. It should also be noted that I can not reject the hypothesis that an increase in *aa* completely offsets an equal increase in school quality.

for black students going to schools with higher levels of affirmative action disappears as the market acquires more information as to their true ability. In 1997, when the graduates have been out of school for four years, the marginal effect of affirmative action on the wages of black graduates is -0.0012 with a standard error of 0.0018. By 2003 the coefficient is actually positive but with a large standard error. A similar relationship holds for the extra return to quality received by black graduates above white graduates. The return to quality observed by white students is significant in 1994 and decreases slightly but remains significant in 1997. In 2003 the point estimate of return to school quality for white students has dropped to about one half of its initial value and is no longer significant. In 1994 the estimated return to school quality for black students is significantly higher than it is for white students, but by 1997 this return is virtually zero. Thus, by 1997 black students no longer realize a wage differential for affirmative action, but they no longer realize positive returns to school quality either.

Also of interest from Table 12 is the coefficient on the percentage of a school population that is black. It is apparent from the actions of universities that diversity of student population is a goal. This view presumably stems from the fact that students are viewed as an input as well as an output in education, and that minorities provide positive externalities in education (Holzer and Neumark, 2000). The coefficient of 0.0046 on *% black* indicates a positive correlation between wages of white graduates and the relative size of black student population as compared to the white student populations. The marginal effect of the size of the black student population on the wages of black graduates is virtually zero. Whether this relationship is due to a causal link or merely a correlation is beyond the scope of this paper, but the finding is consistent with the hypothesis that there are spillovers associated with increasing diversity. Affirmative action may benefit a school's graduates (and society) as a whole, even if it provides no direct benefit to black graduates.

5.1. Small Schools and Measurement Error in Affirmative Action

I do not explicitly restrict the sample used in Table 12 based on number of observations from a school, but the construction of aa requires that all schools must have at least three observations, with a minimum of one black and one white observation. Affirmative action is measured by the difference in average entrance exam percentile and omits the individual of observation, so there must be at least one other black and one other white student observed from the school. Although this measure gives information as to the true level of affirmative action, it is inherently noisy. This is especially true for smaller schools with fewer observations. Table 14 reports the results of a regression of the natural log of 1994 wages on X_1 and X_2 with each column progressively eliminating schools with smaller observed samples. The first column is identical to the 1994 results from Table 12 as this is the implicit restriction created by the construction of aa. The second, third, and fourth columns respectively limit the sample to individuals from schools with at least six, ten, or fourteen observations. Each column gives a progressively more accurate measure of affirmative action, but this comes at the cost of eliminating individuals from smaller schools.

Most of the estimated coefficients reported in Table 14 remain fairly consistent as the sample becomes more restrictive. This is especially true for the coefficient on aa, the variable measuring affirmative action. The estimated marginal effects of aa on black graduates is exactly the same when

TABLE 14

Ordinary Least Squares regression results of the natural log of hourly wages for 1994 on the X1 and X2. Standard errors are reported in parenthesis. The columns are restricted to schools with the given number of black and white observations. Included variables that are not reported are: tenure and its square as well as their interaction with the black indicator, experience and its square as well as their interaction with the black indicator, age and its square, a set on indicator variables for previous degrees (5), a set of indicator variables of subsequent degrees (5), an in-state indicator, a set of indicator variables for region (5), a public school indicator, a set of indicators for the school's Carnegie Classification (4), and a set of indicators for college major (6).

		Dependant Variable: Natural log of 1994 hourly wage			
School Obs. Restriction		# black ≥ 1 # white ≥ 1	# black ≥ 3 # white ≥ 3	# black ≥ 5 # white ≥ 5	# black ≥ 7 # white ≥ 7
University Variables	*aa*	0.0008	0.0007	0.0009	0.0016*
		(0.0006)	(0.0007)	(0.0008)	(0.0009)
	quality	0.0024***	0.0025***	0.0008	0.0028*
		(0.0009)	(0.001)	(0.0013)	(0.0017)
	%black	0.0046***	0.0045***	0.005**	0.0088**
		(0.0015)	(0.0018)	(0.0027)	(0.0039)
Individual Variables	black	0.1784	0.0066	0.131	0.0905
		(0.2377)	(0.2691)	(0.3522)	(0.4883)
	male	0.0412**	0.0433**	0.0233	0.0202
		(0.018)	(0.0198)	(0.0231)	(0.0265)
	GPA	0.0249	0.0242	0.0186	−0.0083
		(0.0202)	(0.0221)	(0.0255)	(0.0296)
	percentile	0.0004	0.0006	0.0006	0.0012**
		(0.004)	(0.0004)	(0.0005)	(0.0005)
Black Interactions	*aa*	−0.0064***	−0.0063**	−0.0053*	−0.0082**
		(0.0024)	(0.0026)	(0.0031)	(0.0042)
	quality	0.0051***	0.0056**	0.0049	0.0101**
		(0.0025)	(0.0027)	(0.0035)	(0.0049)
	%black	−0.0035	−0.0037	−0.005	−0.0008
		(0.0023)	(0.0025)	(0.0032)	(0.0145)
	GPA	0.0061	0.0589	0.0225	−0.0216
		(0.0734)	(0.0854)	(0.1034)	(0.1254)
	percentile	−0.0045***	−0.0047***	−0.0042**	−0.0062***
		(0.0016)	(0.0018)	(0.002)	(0.0023)
	N	1817	1554	1143	756
	R^2	0.19	0.17	0.19	0.26

school sample size is restricted to 6 as when it is not explicitly restricted. It fluctuates a small amount in the more restrictive samples but remains significant and within the standard errors of the less restrictive estimates. The estimated effects of school quality are slightly less consistent. The marginal effects are almost identical for both black and white students in the two least restrictive samples but change in the more restrictive samples.

5.2. Empirical Specification

The results reported in Table 12 are given without controlling for an individual's occupation and industry. This is done because occupation and industry are outcomes that depend on education and might be related to race as well as affirmative action. Looking at the results while controlling for these variables does reveal some information. Specifically, it gives the effects of affirmative action within a particular job. Other included variable might have similar problems. In my data, experience is created from spells of unemployment. If affirmative action makes it harder for black graduates to find jobs then the inclusion of experience will create results that understate the effects of affirmative action. Finally, according to the "fit" hypothesis, black students who attend more selective schools because of affirmative action may choose less lucrative majors. Table 15 reports the results of a regression of the natural log of 1994 hourly wages on X_1 and X_2 progressively omitting the variables of concern.

The first column of Table 15 estimates equation (55) while also including controls for industry and occupation.[63] The second column of Table 15 reports results identical to the 1994 result of Table 12, while column III omits industry, occupation and experience. In general, the measured

[63] Fifteen categories were used to control for industry and twelve categories were used to control for occupation.

TABLE 15

Ordinary Least Squares regression results of the natural log of hourly wages in 1994 on the given variables as well as those listed below. Inclusion of industry, occupation, experience and major is denoted by an X. There are 15 industry categories, 12 occupation categories, 6 major categories, and experience is included along with its square as well as their interaction with the black indicator. Included variables that are not reported are: tenure and its square as well as their interaction with the black indicator, age and its square, a set on indicator variables for previous degrees (5), a set of indicator variables of subsequent degrees (5), an in-state indicator, a set of indicator variables for region (5), a public school indicator, and a set of indicators for the school's Carnegie Classification (4).

		I	II	III	IV
University Variables	*aa*	0.0001	0.0008	0.0008	0.0006
		(0.0005)	(0.0006)	(0.0006)	(0.0006)
	quality	0.0017**	0.0024***	0.0024***	0.0026***
		(0.0009)	(0.0009)	(0.0009)	(0.001)
	%black	0.0043***	0.0047***	0.0046***	0.006***
		(0.0014)	(0.0015)	(0.0015)	(0.0016)
Individual Variables	black	0.2737	0.207	0.0417	0.1435
		(0.2282)	(0.2318)	(0.2258)	(0.2371)
	male	0.031*	0.041**	0.0431**	0.0688***
		(0.0177)	(0.018)	(0.018)	(0.0184)
	GPA	0.0227	0.0247	0.0232	0.0122
		(0.0196)	(0.0202)	(0.0201)	(0.0208)
	percentile	0.0002	0.0035	0.0003	0.0008**
		(0.0004)	(0.004)	(0.0004)	(0.0004)
Black Interactions	*aa*	−0.0061***	−0.0064***	−0.0064***	−0.0072***
		(0.0023)	(0.0024)	(0.023)	(0.0025)
	quality	0.0045**	0.005***	0.0052***	0.0054**
		(0.0024)	(0.0025)	(0.0025)	(0.0026)
	%black	−0.003	−0.0035	−0.0034	−0.0039*
		(0.0022)	(0.0022)	(0.0023)	(0.0024)
	GPA	−0.0342	0.0056	0.0308	0.0098
		(0.0704)	(0.0733)	(0.0718)	(0.0755)
	percentile	−0.0041***	−0.0045***	−0.0043***	−0.0048***
		(0.0016)	(0.0016)	(0.0016)	(0.0017)
Variables Included	ind.&occ.	Y			
	Experience	Y	Y		
	Major	Y	Y	Y	
	N	1813	1817	1817	1817
	R^2	0.27	0.18	0.16	0.07

coefficient on the interaction of the black indicator and affirmative action is fairly robust to the specification of the model. The significant coefficients in columns I and III are virtually identical to those in column II. These results indicate that affirmative action does not push students, black or white, into less lucrative industries or occupations.

The inclusion of industry and occupation does slightly decrease the return to school quality for both black and white graduates, indicating that students from more selective schools enter into more lucrative industries and occupations. Under this specification, there is a one to one trade off between affirmative action and school quality for black students, indicating that their wages depend entirely on the average ability level of other black students regardless of the ability level of white students at their school. The difference between the estimated return to quality in column I and column II indicates that one benefit black graduates gain from going to higher quality schools is the ability to enter into higher paying industries and occupations. The results also suggest that this benefit may account for the 25% of returns to school quality that are not offset by affirmative action. In other words, if a black student is able to "buy" one point of school quality at the price of one point of affirmative action, the affirmative action may completely offset the direct wage gains to school quality, but the student at the higher quality school retains the ability to take a job in a more lucrative industry or occupation. This result makes sense if the quality of employers a graduate is exposed to is correlated with school quality, but employers fully discount for affirmative action.

As shown in column IV of Table 15, the absence of controls for college major does increase size of the marginal effect of affirmative action on black wages form -0.0056 to -0.0062. This lends some support to the hypothesis that black students admitted under affirmative action may struggle in school. These results indicate that under affirmative action, black students may choose less lucrative majors. I proceed maintaining the specification from column II, but it is important to note that affirmative action may have detrimental effects on the choice of major for black students.

5.3. Endogenous Affirmative Action

It might be believed that *aa* is correlated with wage for non-causal reasons. For instance, it could be that students who receive a poor realization on their entrance exam are more likely to take advantage of affirmative action programs in an attempt to better "match" themselves. If this occurs, then OLS estimates will underestimate the effects of affirmative action. A similar result holds if schools have better measures of ability than entrance exams. Alternatively, students receiving a high realization of entrance exam scores might be more likely to take advantage of affirmative action, in which case OLS will overestimate the effects of affirmative action. For these reasons, one might expect the OLS estimates to biased, but it is not clear in which direction. To account for possible bias, I next estimate equation 55 after instrumenting affirmative action.[64] The results are reported in Table 16.

To obtain the results of Table 16 I instrument *aa* and the interaction of *black* and *aa* using the percentage of the school's state that voted Democratic in the 1992 presidential election[65] ($\%dem$) and the interaction of *black* and $\%dem$. There is no reason to believe that a state's political

[64] This procedure will also solve some of the problems associated with the measurement error associated with affirmative action.

[65] The use of alternative instruments such as the percentage of the state voting Democratic in the 1988 presidential election, the number of Democratic senators representing a state in 1992, and whether the individual attended a private high school provide similar results.

TABLE 16
Instrumental Variables regression results of the natural log of hourly wages on the X1 and X2. Standard errors are reported in parenthesis.

	1994	1997	2003
aa	−0.0028	−0.0182*	−0.0127
	(0.0043)	(0.0061)	(0.0081)
black**aa*	−0.0057	−0.0016	−0.1159
	(0.0133)	(0.0237)	(0.1283)
N	1817	2108	2284

makeup should influence wages after controlling for region. On the other hand, it is highly likely that the political persuasion of the state should influence admission polices within the state. In fact, a regression of *aa* on $\%dem$, $black * \%dem$, X_1, and X_2 places a highly significant coefficient on $\%dem$. The coefficient fluctuates between −0.4 and −0.47 depending on the year, with a standard error never above 0.084. Similar results hold for the interaction of *black* and $\%dem$.

As is common with IV estimates, the standard errors in Table 16 are much larger than those of the OLS estimates. Although the large standard errors make inference difficult, the point estimates for the coefficients on *aa* have all become negative. The point estimates of the interaction coefficient for 1994 and 1997 have changed little. Because the coefficient on *aa* in these years has grown, the marginal effects have become more negative for black graduates. These results suggest that if *aa* is endogenous, the OLS estimates likely underestimate the true effect of affirmative action on wages for both black and white graduates.

Although there may be some concern that *aa* is endogenous, the lack of a clear reason for believing a specific direction of bias may ease this concern. Indeed, testing for endogenaity[66] of *aa* using the Durbin-Wu-Hausman test

[66] The second instrument used for this test was the percentage of the school's state

produces a test statistic of 0.91. Comparing this to the critical value taken from a $\chi^2(2)$ distribution, I cannot reject the null hypothesis that aa is exogenous.

6. CONCLUSION

Much economic research has been done on the effects of preferential college admission on educational outcomes. This is important, but economists are often concerned with educational outcomes because of their strong correlation with labor market outcomes. To my knowledge this is the first attempt to directly link college affirmative action policies to the wages of college graduates. College not only serves as an institution for producing human capital, but also as a signal to employers about true ability. Rational employers will take into account the distribution of ability in a given population based on easily observable characteristics such as race. The introduction of affirmative action plans will change this distribution of ability for selected groups.

This paper develops a model that examines the effects of preferential college admission on the wages of college graduates, and then applies the model to data collected by the B&B survey of 1993. I show that affirmative action can potentially hurt minority graduates in two ways. First, it decreases employers expectations about the ability of graduates of a particular university, causing them to make lower wage offers. Second, affirmative action decreases the signaling value of school quality.

Using data from the B&B, I find evidence to support the theory. My results indicate that the initial expected wage of black graduates from a

voting independent in the 1992 presidential election. Recall that Ross Perot ran in 1992 and captured a significance portion of the popular vote.

school with the average level of affirmative action will be approximately 13% lower than those of a black graduate from a school of similar quality with no affirmative action. This wage differential disappears as time passes. There is no significant change in wages for white graduates.

Although I find evidence that affirmative action lowers wages of black college graduates, it is not clear that the policy will actually make them worse off. Affirmative action lowers beneficiaries' wages conditional on school quality, but under affirmative action they are able to go to higher quality schools. I find that the loss in wages faced by black students due to increasing affirmative action offsets approximately 75% of the return to increasing quality. These results suggest that the wages of black graduates depend on the average ability level of other black graduates from the same school, with the distribution of ability in the white population of graduates having little effect.

In this paper I treat affirmative action policies as though they differ only in size, but in reality they may differ in other substantive ways. Preferential admission is instituted through many different types of policies, each of which will have different effects on the student population. Race neutral policies such as the "percent plans" instituted by Texas, California and Florida may actually cause more problems than the explicit affirmative action plans they are meant to replace. Under these plans the variance in true ability might actually increase for both black and white students. They might also lead to lower expected ability for both groups of students. Some researchers such as Fryer, Loury, and Yuret (2003) have begun to look at the implications of these plans but further research is needed into the effects of policy structure on labor market outcomes.

The implications of the model I develop extend beyond its direct application to preferential college admission. For example, the practice of "race norming" used the United States Employment Service (USES) in the 1970s likely created similar results. The USES used an aptitude test to refer the best applicants to jobs. Because blacks scored lower on the aptitude test, the USES normed the results by race (Jencks, 1998). Given their knowledge of this practice, it is likely that employers held lower expectations of black referrals. In a situation such as this, the practice could result in lower wages for blacks, or alternatively, fewer black hires if the firms cannot pay differentially because of regulation or because they are only willing to offer the minimum wage.

The model also extends beyond the labor market. Similar results will arise anytime one institution or individual screens quality for a second. Take for instance a used car lot that sells two types of used cars, "certified" used cars that have been checked by a mechanic and used cars being sold with no such quality screening process.[67] The cars may initially come from the same population, and are sold on the same lot, yet the certified used cars will command a higher price because the initial screening process informs the consumer that the cars are less likely to break down or require repair.

[67] For this example I am assuming that cars that fail the screening are not sold as unscreened cars.

REFERENCES

[1] Aigner, D., and G. Cain, "Statistical Theories of Discrimination in the Labor Market." *Industrial and Labor Relations Review,* 30 (January 1977), 175-187.

[2] Altonji, J., and C. Pierret,"Employer Learning and Statistical Discrimination." *Quarterly Journal of Economics*, 116 (February 2001), 313-350.

[3] Becker, G. *The Economics of Discrimination.* 1st ed. Chicago: University of Chicago Press, 1957.

[4] Behrman, Jere R., and Mark R. Rosenzweig. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *The American Economic Review*, March 2002, vol. 92, no. 1: 323-334.

[5] Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. "Why the Apple Doesn't Fall Far: Understanding the Inter-generational Transmission of Human Capital." *The American Economic Review*, March 2005, vol. 95 no. 1: 437-449.

[6] Bowen, W.G., and D. Bok, *The shape of the river: Long term consequences of considering race in college and university admissions,* Princeton, NJ, Princeton University Press.

[7] Bureau of Labor Statistics (BLS) *NLS: NLSY79 User's Guide 2002*, Washington, D.C.: U.S. Department of Labor, 2002.

[8] Bureau of Labor Statistics (BLS) *NLSY: Attachment 106: Profiles*, Washington, D.C.: U.S. Department of Labor.

[9] Card, D., and A. Krueger, "Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas," NBER Working Paper 10366 (2004).

[10] Carneiro, P., J. Heckman and D. Masterov, "Labor Market Discrimination and Racial Differences in Premarket Factors." *Journal of Law and Economics,* 48 (April 2005), 1-39.

[11] Chan, J., and E. Eyester, "Does Banning Affirmative Action Lower College Student Quality?" *American Economic Review,* 93:3 (2003), 858-72.

[12] Chevalier, Arnaud. "Parental Education and Child's Education: A Natural Experiment." IZA Discussion Paper No. 1153, May 2004.

[13] Coate, S., and G. Loury, "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review,* 83:5 (1993), 1120-40.

[14] Currie, Janet, and Enrico Moretti. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence From College Openings and Longitudinal Data." *Quarterly Journal of Economics*, November 2003, vol. 118 no.4:1495-1532.

[15] Datcher Loury, L., and D. Garman, "Affirmative Action in Higher Education," *American Economic Review,* 83:2 (1993), 99-103.

[16] ––– "College Selectivity and Earnings," *Journal of Labor Economics,* 13:2 (1995), 289-308.

[17] Dickson, L., "Does ending affirmative action in college admissions lower the percent of minority students applying to college?" Working Paper, (2004).

[18] Farber, H. and R. Gibbons, "Learning and Wage Dynamics." *Quarterly Journal of Economics,* 111 (1996), 1007-1047.

[19] Fryer, R., G. Loury, and T. Yuret, "Color-Blind Affirmative Action," NBER Working Paper 10103 (2003).

[20] Green, W., *Econometric Analysis.* 5th ed. Upper Saddle River, New Jersey: Prentice Hall, 2003.

[21] Holzer, H., and D. Nermark, " Assessing Affirmative Action," *Journal of Economic Literature,* 38:3 (2000) 483-568.

[22] Jencks, C., "Racial Bias in Testing," In C. Jencks and M. Philips eds. *The Black-White Test Score Gap.* Washington, D.C.: Brookings Institution, 1998, 55-84.

[23] Kane, T., "Racial and Ethnic preference in college admissions," In C. Jencks and M. Philips eds. *The Black-White Test Score Gap.* Washington, D.C.: Brookings Institution, 1998, 431-56.

[24] Lang, K. and M. Manove, "Education and Labor-Market Discrimination" Unpublished manuscript. Boston: Boston University, February 2006.

[25] Long, C. "College Applications and the Effects of Affirmative Action," *Journal of Econometrics*, 121 (2004), 319-342.

[26] Lundberg, S. and R. Startz, "Private Discrimination and Social Intervention is Competitive Labor Markets." *American Economic Review,* 73 (1983), 340-347.

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

[28] Oettinger, G., " Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap." *Journal of Labor Economics,* 14 (1996), 52-78.

[29] Oreopoulos, Philip, Marianne E. Page, and Ann Huff Stevens. "Does Human Capital Transfer From Parent to Child? The Intergenerational Effects of Compulsory Schooling." *Journal of Labor Economics*, October 2006, vol. 24 no. 4: 229-760.

[30] Plug, Erik. "How do Parents Raise the Educational Attainment of Future Generations?" IZA Discussion Paper No. 652, November 2002.

[31] Rosenzweig, Mark, and Kenneth I. Wolpin. "Are There Increasing Returns to the Inter-generational Production of Human Capital? Maternal Schooling and Child Intellectual Achievement." *The Journal of Human Resources,* Spring 1994, vol. 29, no. 2: 670-693.

[32] Sacerdote, Bruce. "The Nature and Nurture of Economic Outcomes." *The American Economic Review*, May 2002, vol. 92, no. 2: 344-348.

[33] Vars, F., and W. Bowen, "Scholastic Aptitude, Test Scores, Race, and Academic Performance in Selective Colleges and Universities," In C. Jencks and M. Philips eds. *The Black-White Test Score Gap.* Washington, D.C.: Brookings Institution, 1998, 457-79.

[34] Wigdor, A. and B. Green Jr., eds. *Performance Assessment for the Workplace.* 2 vols. Washington: National Academy Press, 1991.

[35] Wooldridge, Jeffrey W. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Massachusetts: The MIT Press, 2002.

VITA

Christopher Laurence Bartlett was born in El Paso, Texas on June 15, 1978, the son of Laurence W. and Deborah Bartlett. After graduating from J. M. Hanks High School, El Paso, Texas, in 1997, he enrolled at the University of Oklahoma. At the University of Oklahoma, he studied economics, management, and math, and was a member of the men's gymnastics team. He received a Bachelor of Business Administration from the University of Oklahoma in May 2002. In September 2002 he entered the Graduate School of the University of Texas and was awarded a Master of Science in Economics in May of 2004.

Permanent Address: 8513 Tangleridge Drive, Fort Worth, Texas, 76123

This dissertation was typed by the author.