Persistent Racial Wage Inequality

Kate Antonovics*

Department of Economics

University of California–San Diego

9500 Gilman Drive, 0508

La Jolla, CA 92093-0508

kantonov@ucsd.edu

(858) 534-2973

October 12, 2002

Abstract

This paper attempts to understand the forces that have lead to persistent racial wage inequality by developing a dynamic model of statistical discrimination that accounts for the transmission of earnings across generations. The parameters of this model are then estimated using data from the 1970 and 1990 U.S. Census. The results indicate that racial disparities in the quality of information that firms receive about worker productivity are the primary cause of racial wage inequality in 1990. The results also indicate that neither the persistence of income across generations nor the presence of coordination failures explains a sizable fraction of ongoing inequality.

JEL Classification: J70, J15, D82, J62

Keywords: Discrimination, Intergenerational Income Mobility

*I would like to thank John Kennan, Derek Neal, Peter Norman, Arthur Goldberger, Larry Samuelson, Yuichi Kitamura and Andrea Moro for their help on this project. I am also greatful to Lianjun Su who provided excellent research assistance.

1 Introduction

Racial wage inequality has been a persistent feature of the U.S. labor market. Abstracting from cultural and genetic explanations for this ongoing inequality, there are two primary reasons why the racial wage gap may have persisted. The first is ongoing racial discrimination and the second is the positive correlation in earnings across generations. While both of these phenomena may be responsible for lasting racial wage inequality, evaluating the relative importance of each may be difficult, particularly if racial discrimination and past inequality are interrelated.

Thus, as a first step towards understanding the sources of persistent inequality, this paper develops a dynamic model of statistical discrimination that accounts for the correlation in earnings across generations. Then, using data from the 1970 and 1990 U.S. Census, the model's fundamental parameters are estimated in order to evaluate the sources of racial wage inequality in 1990.

In the model, income is transmitted across generations through parental investments in the human capital of children, and discrimination arises because firms are not able to perfectly observe these parental investments. Thus, firms use a noisy signal of worker productivity and their prior beliefs about the average productivity of workers from different racial groups in order to form their assessment of each worker's expected marginal product. In equilibrium, parents invest at the exact rate postulated by firms so that the firms' beliefs are self-confirming. Within this context, there are two sources of discrimination. The first is the existence of multiple equilibria. Thus, even if two groups are identical ex ante, coordination failures alone may lead the firms' equilibrium priors to differ across racial groups. The second source of discrimination is past inequality and the fact that parents from disadvantaged racial groups have fewer financial resources with which to invest in their children's human capital. As a result, on average firms will rationally judge workers from the disadvantaged group to be less qualified.

The model described above is estimated using weekly wage data for blacks, whites and Hispanics from the 1970 and 1990 U.S. Census. In addition, the empirical specification allows for a third source of racial wage inequality. In particular, the precision with which workers are able to signal their true productivity is allowed to vary across racial groups. Since the ability of workers to signal their productivity will affect parents' investment incentives, this asymmetry can also lead to racial wage inequality and racial wage discrimination. The parameter estimates are used to simulate what wages would have been if all three racial groups had the same signaling technology, the same distribution of income in the parent's generation, and if coordination failures were eliminated. These simulations allow the relative importance of these three sources of inequality to be evaluated.

The crucial feature that allows estimation of this model's fundamental parameters is the fact that, conditional on parental investment decisions, the equilibrium wage distribution is uniquely determined by a subset of the model's parameters. Thus, once functional forms have been adopted, it is possible to estimate that subset directly. The remaining parameters can then be estimated in an additional stage[1]. Here, estimation proceeds in three stages. In the first stage, the wage distribution for each racial group is adjusted to account for racial differences in college graduation rates. In the second stage, a subset of the model's parameters are estimated by minimizing the distance between the quantiles of the wage distribution that are predicted by the model and those that are observed in the data. The remaining parameters, those governing the distribution of parental investment costs, are uncovered in a third stage by linking the second stage estimates from 1970 with those from 1990 through the parental investment decision.

According to the model, the parameter estimates suggest that neither coordination failures nor the persistence of earnings across generations is the primary source of racial wage inequality in 1990. Rather, racial wage inequality is found to be primarily the result of racial differences in workers' ability to signal their true productivity. In particular, it is found that black and Hispanic workers find it more difficult to signal their true productivity than whites, and, since this lowers the incentives for parents to invest in their children's human capital, blacks and Hispanics acquire less human capital and subsequently earn lower wages than whites. Moreover, the fact that these same signaling differentials are found to exist in 1970 suggests that informational asymmetries may play an important role in explaining ongoing racial wage inequality.

[1]This technique was first employed by Moro (2001)

The paper proceeds as follows. Section 2 discusses related research. Section 3 presents the model and Section 4 discusses the equilibria of the model. Section 5 presents the estimation technique and Sections 6 and 7 discuss the data and the results. Finally, Section 8 concludes.

2 Related Research

The model presented in this paper builds upon the theory of statistical discrimination pioneered by Arrow(1972, 1973), McCall(1972) and Phelps(1972). In models of statistical discrimination, discrimination is explained as a rational response to uncertainty in labor markets. In these models, workers differ by some observable characteristic (i.e. the color of their skin) and can be classified into groups based on that characteristic. Firms cannot perfectly observe worker productivity and base their assessment on a noisy signal and on prior beliefs about the productivity of workers in different groups. Within this framework, discriminatory outcomes can occur either in the presence of exogenous group differences (for example, in the amount of noise associated with each group's productivity signal) or in the presence of multiple equilibria.

Among the papers on statistical discrimination, this paper is related to Coate and Loury (1993) in that the presence of multiple equilibria is one of the major sources of potential racial wage inequality. However, this paper also shares elements of Lundberg and Startz (1983) since firms may receive more reliable productivity signals for workers from one racial group than they do for workers from some other group. In addition, similar to Lundberg and Startz, this informational asymmetry may lead to differences in workers' incentives to invest in human capital.

Nonetheless, the model differs from previous models of statistical discrimination in two significant ways. First, the model developed here is dynamic and considers the evolution of racial discrimination over many generations of workers. Thus, it is uniquely well-suited for examining the persistence of racial wage inequality. Second, in previous models of statistical discrimination, uncertainty leads to discrimination either because of a coordination failure or because of exogenous group differences. However, this paper also considers how these forces can lead to *endogenous* group differences (in this case, racial wage differentials) that, in and of themselves, lead to future discrimination.

This paper also builds upon a large body of theoretical literature on intergenerational income mobility, and a salient feature of the model is the positive association between parental income and investment in children. According to standard theories of intergenerational income mobility, parents find it optimal to invest in their children's human capital until the rate of return on these investments is equal to the rate of return on other investments (i.e. the interest rate). If capital markets are perfect, poor families can borrow against their children's future earnings in order to finance investment, and thus, there is no relationship between parental resources and the human capital of children. However, it is often argued that parents may not have perfect access to capital markets because it may be difficult to enforce contracts that require children to repay debts that were incurred by their parents. Loury(1981) and Becker and Tomes(1986) examine how standard models change when capital markets fail. In their models, parents must finance investments in children through reductions in their own consumption. Thus, higher levels of parental income are associated with higher levels of human capital investment in children. In the model in this paper, poor parents are assumed to have higher investment costs than wealthy parents, and, thus, they are less likely to invest in their children. The interpretation is that imperfect capital markets raise investment costs for less wealthy parents.

On the empirical side, a number of previous papers have attempted to identify the forces that have led to changes in the magnitude the racial wage gap over time. For example, Heckman and Donohue (1991), Card and Lemieux (1996) and Moro (2001) attempt to explain movement in the black-white wage gap over various periods in American history. This paper differs from these studies in that it focuses on examining why the racial wage gap has persisted rather than on the forces that have led to changes in that gap.

Of the papers that examine changes in racial wage inequality over time, the theoretical and empirical approach of this paper is most closely related to Moro (2001). Using a similar model to the one presented in this paper, Moro attempts to evaluate whether or not changes in the black-white wage gap over time are the result of economy-wide shifts between different equilibria. This paper differs from Moro's in that it considers the effect of past inequality on current racial wage differentials. As such, it is well suited for examining the persistence of racial wage differentials. Nonetheless, the fundamental insight that allows the parameters of this class of models to be estimated was first noted by Moro.

Finally, in keeping with the model, the empirical section of this paper takes as given the fact that racial wage inequality is at least partly the result of uncertainty about worker productivity. A natural question to then ask is whether there is any evidence that uncertainty contributes to racial wage differentials. Unfortunately, there have been very few studies that examine this issue, and those that do offer conflicting evidence. Oettinger (1996) builds a model of statistical discrimination in which employers cannot perfectly observe match quality, and although the mean productivity of job matches is the same across racial groups, employers receive noisier signals about the match productivity for blacks than whites. According to his model, if employers eventually learn the true match productivity, then there should exist no black/white wage gap at labor force entry, but one should emerge with experience since blacks receive smaller gains from job mobility. Altonji and Pierret (1997) also develop a model of statistical discrimination. However in their model employers receive comparable signals about the productivity of blacks and whites, but blacks have a lower mean productivity level than whites. In contrast to Oettinger's theoretical predictions, in their model, there should exist a black-white gap at labor market entry and, moreover, there should be no change in this gap over time. Thus, while Oettinger interprets the fact that the black-white wage gap grows with experience as evidence in favor of statistical discrimination, Altonji and Pierret interpret this trend as evidence against statistical discrimination. In addition neither of these papers account for the fact that statistical discrimination may lead to racial differences in on-the-job training and that it is this phenomenon that is responsible for the observed differences in the wage profiles of blacks and whites. In sum, the empirical relevance of labor market uncertainty in explaining racial wage differentials is still unknown.

3 Model

Individual Behavior

In the model, individuals are organized into families consisting of one parent and one child. The time horizon is infinite. However, individuals live for only two periods. In the first period, they are children and do not participate in the labor force. At the beginning of the second period, they become parents and receive income by selling their leisure in the labor market. As parents, individuals value their own consumption and the wage their child will earn as an adult. Parents can influence their child's wage by investing in their child's human capital. Parental investments can take the form of both market goods and time, and might include things like music lessons, tutoring, buying a house in a good school district or sending their child to private school. While these investments enhance productivity, they are not easily observable by firms. For simplicity, investment is modeled as a binary choice. Let the expected utility of a parent from group j at time $t-1$ be given by:

$$E_{t-1}^{j}[U] = w_{t-1} + E_{t-1}^{j}[w_t|\delta] - \delta c$$

where w_{t-1} is the parent's wage, w_t is the child's wage, c is the cost of investment and δ is an indicator variable that equals one if the parent invests and zero otherwise.

A worker whose parent invests is "qualified", otherwise, the worker is "unqualified". When workers enter the labor market, they receive a productivity signal $\theta \in [0,1]$. For workers who are qualified, θ is drawn from the p.d.f. $f_q(\cdot)$, and for those who are unqualified, θ is drawn from the p.d.f. $f_u(\cdot)$. It is assumed that $\frac{f_q(\theta)}{f_u(\theta)}$ is strictly increasing in θ. Since this strict Monotone Likelihood Ratio Property implies that higher values of the productivity signal are more likely if a worker is qualified, workers with higher productivity signals earn higher wages. Thus, since parents care about their child's future wage, they have an incentive to invest in their children.

However, parents must balance this incentive to invest with the cost of investment, c. Costs are assumed to be distributed on $[\underline{c}, \overline{c}]$ where $\overline{c} > \underline{c} > 0$. Let $G(c|w)$ denote the distribution of investment costs conditional on a parent's wage and let $g(c|w)$ denote the corresponding density. It is assumed that high-wage parents have lower investment costs than low-wage parents in the sense of first order stochastic dominance. That is, $G(c|w')$ first order stochastically dominates $G(c|w)$ for all $w' < w$. The interpretation is that low-wage parents face borrowing constraints[2].

Finally, individuals belong to one of two identifiable groups, blacks and whites, denoted group b and group w respectively.

[2] If investment is composed of time and market goods, then this assumption implies that the cost of investment is dominated by the cost of market goods since the cost of time will be lower for low-wage parents.

Firm Behavior

It is assumed that there are two identical, risk-neutral firms that compete for workers and that exist for only a single period. At each firm, qualified workers are more productive than unqualified workers. Qualified workers produce y_q and unqualified workers only produce y_u, where $y_u < y_q$.

Firms do not know whether a worker's parent invested. Instead, they only observe the worker's racial identity and the productivity signal. Based on these two pieces of information and the firm's prior beliefs about the probability that workers from each group are qualified (i.e. that parents invested), firms form posterior beliefs about the probability that members from each group are qualified. Let π_t^j denote a firm's prior belief that a worker from group j at time t is qualified, and let $p(\theta, \pi_t^j)$ denote a firm's posterior belief that a worker with signal θ is qualified. Then

$$p(\theta, \pi_t^j) = \frac{\pi_t^j f_q(\theta)}{\pi_t^j f_q(\theta) + (1 - \pi_t^j) f_u(\theta)}. \tag{1}$$

The Monotone Likelihood Ratio Property guarantees that a firm's posterior belief that a worker is qualified is increasing in θ.

Timing of the Within-Generation Game

Taking the distribution of wages in the parents' generation as given, the timing of the game within a single generation of workers can be described as follows:

Stage 1. Parents have children, learn their investment cost and decide whether to invest in their child's human capital. The distribution from which a parent's investment cost is drawn will depend on his or her wage.

Stage 2. At the beginning of the next period, children become workers and enter the labor market. When they do so, they receive a signal $\theta \in [0, 1]$. If the worker's parent invested, then the signal θ is distributed according to the p.d.f. f_q. Otherwise, it is distributed according to the p.d.f. f_u.

Stage 3. Firms compete for workers by simultaneously announcing wage schedules. Wages are allowed to depend on the signal and group identity, so a pure action of firm i at time t is a pair of functions $w^j_{t,i} : [0,1] \to \Re_+$, $j = b, w$.

Stage 4. Workers observe wage schedules and decide which firm to work for.

This "stage game" or "within-generation game" is then repeated over an infinite number of generations. In contrast to a true repeated game, however, these within-generation games are linked over time since the equilibrium wage distribution in one generation will affect parental investment costs in the subsequent generation. Nonetheless, it is convenient to first discuss the equilibrium of the within-generation game. It is then straightforward to characterize the equilibrium of the infinite-period model.

4 Analysis of the Model

Within-Generation Equilibrium

The equilibria of the within-generation game are found through backwards induction. In Stage 4, workers decide which firm to work for. A worker's best response is to work for the firm that offers him or her the highest wage.

In Stage 3, firms announce wage schedules. Suppose that, in equilibrium, workers are paid their expected marginal product. Thus, given the probability that workers from each group invest, π^j_t, the wage for a worker from group j at time t with productivity signal θ would be given by:

$$w(\theta, \pi^j_t) = p(\theta, \pi^j_t)y_q + (1 - p(\theta, \pi^j_t))y_u. \tag{2}$$

Formally, a firm's action in Stage 3 (the wage offer schedule) is only a function of a worker's productivity signal. However, in equilibrium a firm's best response will depend on the firm's prior beliefs, and it is convenient to explicitly note this dependence by writing wages as a function of π^j_t. Following the Bertrand price competition logic, it can be shown that firms are playing best responses if and only if they choose wage schedules according to $w(\theta, \pi_t^j)$. The following proposition states this formally.

Proposition 1 *Let the fraction of investors from group j at time t be given by π_t^j. Then firms are playing best responses if and only if they choose wage schedules according to $w(\theta, \pi_t^j)$.*

The proof can be found in the appendix.

In Stage 2 children receive productivity signals. Finally, in Stage 1, parents make investment decisions. Let $B(\pi_t^j)$ denote the benefit of investing in a child from group j who will enter the labor market at time t when the firms' prior beliefs are π_t^j. The benefit to investing is simply the difference between the child's expected wage if the parent invests and the child's expected wage if the parent does not invest. This is given by:

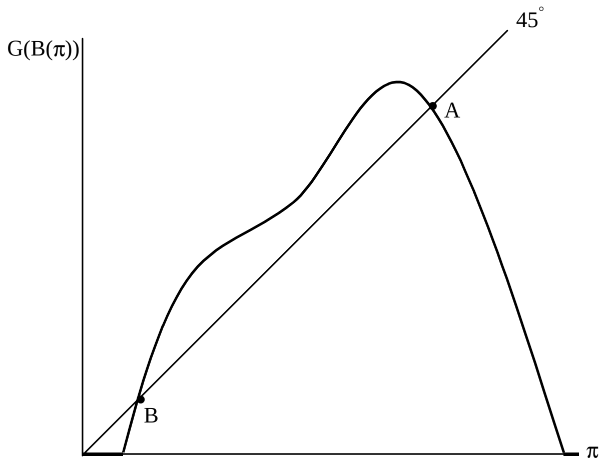
$$B(\pi_t^j) = \int_\theta w(\theta, \pi_t^j)(f_q(\theta) - f_u(\theta))d\theta.$$

Parents play a best response if and only if they invest whenever the cost of investment is less than the benefit of investing. Thus, the probability that optimizing parents at time $t-1$ invest in their child's human capital is

$$G_{t-1}^j(B(\pi_t^j)) = \int_w \int_{\underline{c}}^{B(\pi_t^j)} g(c|w) x_{t-1}^j(w) dc dw. \tag{3}$$

where $x_{t-1}^j(w)$ is the p.d.f. of wages in the parents' generation and where the superscript j in $G_{t-1}^j(\cdot)$ refers to the fact that the distribution of wages may vary across groups in any generation.

Figure 1 depicts the general pattern of how the probability of investment changes as π_t^j changes. Note that the behavior of $B(\pi_t^j)$ drives investment decisions. When firms' prior beliefs are zero, the benefit to investing is zero since all workers earn y_u regardless of their productivity signal. Likewise, when the firms' prior beliefs are equal to one, the benefit to investing is zero since all workers earn y_q. In addition, it is possible to show that when π_t^j is small (close to zero), the benefit to investment is increasing in π_t^j. On the other hand, when π_t^j is large (close to one) the benefit to investment is decreasing in π_t^j.


G(B(π))
45°
A
B
π

I define an equilibrium of the within-generation game to be one in which parents are playing a best response to the wage schedules set by firms given their investment costs, and in which firms are playing a best response to the distribution of parental strategies given their beliefs. Such an equilibrium occurs where the wage schedules are such that the firms' beliefs are self-confirming. In other words, based on their prior belief about the probability that parents from a given group invest, the labor market standard set by firms must be such that parents from each group optimally invest at the rate postulated by firms. Thus, a self-confirming equilibrium in this economy is any π_t^j such that

$$\pi_t^j = G_{t-1}^j(B(\pi_t^j)). \tag{4}$$

The existence of a within-generation equilibrium reduces to the existence of a fixed point in equation (4). A trivial equilibrium always exists at $\pi_t^j = 0$. To see this, note that when $\pi_t^j = 0$, $B(0) = 0$. Since no parents will invest when the benefit to investing is zero, we know that $G(0) = 0$. The following proposition states the conditions under which there exist non-trivial equilibria.

Proposition 2 *If $f_q(\cdot)$ and $f_u(\cdot)$ are continuous on $[0,1]$, $g(\cdot|w)$ is continuous on $[\underline{c}, \overline{c}]$ for all w, and if $G_{t-1}^j(B(\eta)) > \eta$ for some $\eta \in (0,1)$ then there exist multiple, non-trivial within-generation equilibrium for group j at time t.*

The proof is in the appendix.

Infinite-Generation Equilibrium

In any equilibrium of the infinite-generation model, both individuals and firms behave exactly as they do in the within-generation game. To see this, note that since individuals live for only two periods and since adults care only about the wages that their children earn, the decision about whether or not to invest is made without regard to the welfare of any future generations. Similarly, firms are assumed to live for only one period. Thus, in any equilibrium of the infinite-generation model, equation (4) must hold in each period.

Sources of Discrimination

Within this model there are three basic forces that can lead the equilibrium outcome to differ across racial groups.

First, if the distribution of wages in the parents' generation differs for blacks and whites, then so will the distribution of investment costs. In this case $G_{t-1}^{b}(\cdot) \neq G_{t-1}^{w}(\cdot)$. Thus, firms' equilibrium priors will differ across racial groups, and individuals who have identical productivity signals but who are from different racial groups will earn different wages.

Second, even if the distribution of wages is identical across groups, discrimination can arise due to a coordination failure. This can occur whenever there are multiple solutions to equation (4). In Figure 1, for example, there is a self-confirming equilibrium both at point A and point B. The equilibrium at point B will be referred to as a coordination failure since both workers and firms are at least as well off at point A as at point B.

Finally, although not explicitly discussed above, racial differences in the distribution of productivity signals for qualified and unqualified workers can lead to racial wage differentials. In particular, the noisier are these signals, the lower are the incentives for parents to invest in human capital. Thus, if groups cannot effectively signal their qualifications to firms they will invest less in human capital and, as a result, earn lower wages.

The goal of this paper is to empirically determine which of these three factors is primarily responsible for observed racial wage inequality.

5 Estimation

General Discussion

The crucial feature of this model that allows estimation of the model's fundamental parameters is the fact that, conditional on parental investment decisions, the equilibrium wage distribution is uniquely determined by a subset of the model's parameters. Thus, as pointed out by Moro (2001), it is possible to estimate that subset directly. The remaining parameters can then be estimated in an additional stage.

A convenient feature of the model is that linked parent-child data are not needed to obtain parameter estimates. Instead, data on the distribution of wages in two successive generations are sufficient to identify all of the fundamental parameters. This turns out to be critical since almost every data set that contains good earnings data on both parents and children contains very few observations on non-whites. In fact, most previous studies of intergenerational income mobility in the United States use samples that contain virtually no non-white observations. In this paper, I use 1970 and 1990 U.S. Census data on white, black and Hispanic men to obtain data on the distribution of wages. The large sample sizes in these data sets enable me to get precise estimates of the underlying parameters.

In order to estimate the model, functional forms need to be adopted for the productivity signal distributions and for the distribution of parental investment costs. It is assumed that the productivity signal distributions for qualified and unqualified workers are as follows:

$$f_q(\theta) = \gamma_q^j \theta^{\gamma_q^j - 1}$$

$$f_u(\theta) = \gamma_u^j (1-\theta)^{\gamma_u^j - 1}$$

where $\theta \in [0,1]$. These functional forms satisfy the Monotone Likelihood Ratio property for any value of γ_u^j and γ_q^j such that $\gamma_u^j \geq 1$ and $\gamma_q^j \geq 1$. In addition, in order to allow for the possibility that the productivity signal θ may be more noisy for some racial groups than it is for others, the distributions are allowed to vary across racial groups. In particular, as γ_u^j and γ_q^j approach infinity, the variance of both distributions approaches zero. In this case, the distribution of productivity signals for qualified workers would be degenerate at one and the distribution of signals for unqualified workers would be degenerate at zero. Thus, the higher are γ_u^j and γ_q^j, the greater the informational content of the signal θ.

A simple distribution that allows parental investment costs to vary by wage is the following: $G(c|w) = \frac{-m_1}{w} + m_2 c$, where $c \in [\frac{m_1}{m_2 w}, \frac{w+m_1}{m_2 w}]$. In this case, investment costs are uniformly distributed and, as long as $m_1 > 0$, average investment costs rise as wages fall[3]. In addition, recall that the interpretation of the relationship between investments costs and wages is that poor families face borrowing contraints. Thus, this functional form also captures the idea that the relationship between wages and average investment costs are the strongest amongst

[3]The restriction that $m1 > 0$ is not imposed in estimation

relatively poor familes. In particular, for large values of w, changes in w have almost no effect on the distribution of investment costs. In addition, as $m1$ increases, the effect of wages on average investment costs increases. Finally, since parents invest whenever the cost of investment is less than the benefit of investing, m_2 determines the marginal effect of changes in the benefit of investing on invesment.

Estimation proceeds in three stages. These are outlined below.

Three-Step Estimation Strategy

Stage 1: The Residual Wage Distribution

In the first stage of estimation, racial wage differences that result from differences in age and college completion rates are accounted for by regressing wages on a constant, age, age squared and an indicator of whether the individual completed college. The sum of the constant, plus the contribution to wages of being age 40, plus the residual from that regression is referred to as an individual's "residual wage". It represents an individual's wage at age 40 after the returns to college have been eliminated.[4] The remaining variation in wages is taken to be the result of variation in each worker's productivity signal as well as variation (by race) in the firms' prior beliefs about worker productivity.

This procedure of accounting for the returns to college assumes that those returns are independent of the returns to the unobservable human capital investment made by parents. This implies that the cost of going to college is independent of the cost of making the unobservable investment. In addition, it also implies that the unobservable investment and the (observable) college education are not complementary in production. Taken together, these assumptions imply that the returns to college are due purely to productivity gains and that college education provides no information to firms about a worker's unobservable skill level. The alternative extreme assumption would be that college is only a signal of worker skill and that actual productivity is not enhanced at all by going to college. In this case, it would not be necessary to adjust the wage distributions for racial differences in college completion rates

[4] If $\tilde{w}$ is an individual's residual wage, then $\tilde{w} = \hat{\beta}_0 + \hat{\beta}_1(40) + \hat{\beta}_2(1600) + \hat{u}$ where $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ are the estimated coefficients on the constant(=1), age and age squared and where $\hat{u}$ is the residual.

since racial differences in the distribution of productivity signals are already captured by the parameters of the model. As it turns out, none of the model's qualitative results are sensitive to whether the wage distribution is adjusted for racial differences in college completion rates.

Stage 2: The Quantiles of the Residual Wage Distribution

Given the functional forms adopted for the productivity signal distributions, the firms' posterior belief from equation (1) can be rewritten as

$$p(\theta, \pi_t^j) = \frac{\pi_t^j \gamma_q^j \theta^{\gamma_q^j - 1}}{\pi_t^j \gamma_q^j \theta^{\gamma_q^j - 1} + (1 - \pi_t^j)\gamma_u^j (1-\theta)^{\gamma_u^j - 1}}. \tag{5}$$

Since the firm's posterior belief is increasing in θ, and since $y_q > y_u$, we know that wages are also increasing in θ. This monotonic relationship between wages and productivity signals implies that the quantiles of the distribution of productivity signals correspond to the quantiles of the distribution of wages. Put differently, the l^{th} percentile of the distribution of residual wages is equal to $w(\theta, \pi_t^j)$ evaluated at the l^{th} percentile of the distribution of productivity signals. Thus, as a first step towards expressing the quantiles of the residual wage distribution as a function of the underlying parameters, I calculate the quantiles of the distribution of productivity signals as a function of the underlying parameters.

For any number l on the interval $(0,1)$, let $\theta_l(\pi_t^j, \gamma_q^j, \gamma_u^j)$ denote the l^{th} percentile of the distribution of productivity signals as a function of the underlying parameters π_t^j, γ_q^j and γ_u^j. This l^{th} percentile is defined by the value of θ that solves

$$l = F_{\pi_t^j}(\theta) \tag{6}$$

where $F_{\pi_t^j}(\theta) = \pi_t^j \theta^{\gamma_q^j} + (1-\pi_t^j)(1-(1-\theta)^{\gamma_u^j})$ is the cumulative distribution of productivity signals in equilibrium. The fact that this distribution function is strictly increasing guarantees that the equation has a unique root.

As discussed above, the l^{th} percentile of the residual wage distribution for group j at time t is equal to $w(\theta, \pi_t^j)$ evaluated at the l^{th} percentile of the distribution of productivity signals. Thus, by substituting $\theta_l(\pi_t^j, \gamma_q^j, \gamma_u^j)$ into (2), it is possible to write the quantiles of the residual wage distribution for group j at time t as functions of the parameter vector $\beta = \{y_q, y_u, \pi_t^j, \gamma_q^j, \gamma_u^j\}$. In particular, for any $l \in (0,1)$,

$$w_{t,l}^j(\beta) = p(\theta_l(\pi_t^j, \gamma_q^j, \gamma_u^j), \pi_t^j)y_q + (1 - p(\theta_l(\pi_t^j, \gamma_q^j, \gamma_u^j))y_u. \tag{7}$$

The estimation technique that I employ simply finds the value of β that most closely matches these predicted population quantiles with observed sample quantiles. Let $\widehat{\mathbf{w}}_{\mathbf{t}}^{\mathbf{j}}$ denote a $L \times 1$ vector of sample quantiles for group j at time t and let $\mathbf{w}_{\mathbf{t}}^{\mathbf{j}}(\beta_0)$ denote the corresponding vector of predicted population quantiles at the true parameter values. We know that $\sqrt{n}(\widehat{\mathbf{w}}_{\mathbf{t}}^{\mathbf{j}} - \mathbf{w}_{\mathbf{t}}^{\mathbf{j}}(\beta_0)) \rightarrow_d N(0, \mathbf{\Omega}(\beta_{\mathbf{0}}))$ where $\mathbf{\Omega}(\beta_0)$ is the variance-covariance matrix of the sample quantiles (Cramer, 1951, p.368-369).

The estimator of β_0 is the least squares vector $\widehat{\beta}$ where

$$\widehat{\beta} = \arg\min_{\beta}(\widehat{\mathbf{w}}_{\mathbf{t}}^{\mathbf{j}} - \mathbf{w}_{\mathbf{t}}^{\mathbf{j}}(\beta))'(\widehat{\mathbf{w}}_{\mathbf{t}}^{\mathbf{j}} - \mathbf{w}_{\mathbf{t}}^{\mathbf{j}}(\beta)). \tag{8}$$

Chamberlain (1994) shows that this minimum distance estimator is a consistent estimator of β_0. This technique of matching quantiles has also been used by Epple and Sieg (1999).

Since the minimum wage an individual can earn is y_u, and the maximum wage they can earn is y_q, in order to simplify estimation, in both 1970 and 1990, y_u is set equal to the 3rd percentile of the residual wage distribution, and y_q is set equal to the 97th percentile of the residual wage distribution. Thus, in 1970 $y_u = \$82.15$ and $y_q = \$1,543.14$, and in 1990 $y_u = \$49.79$ and $y_q = \$1,804.41$. y_u and y_q are selected in this way to prevent outliers from affecting the parameter estimates. The model's qualitative results are not sensitive to this procedure. Note also that these productivity parameters are restricted to be the same across racial groups.

Because there is reason to believe that there may have been changes in the signal distributions over time, β_0 is estimated separately for 1970 and 1990.

Stage 3: The Distribution of Costs

Given the estimates for y_q, y_u, π^j, γ_q^j and γ_u^j for 1970 and 1990 that were obtained in the first stage, the only parameters remaining to be uncovered are those governing the distribution of investment costs. Since the distribution of investment costs depends on the distribution of total wages in the parent's generation, it is first necessary to uncover the distribution of total wages in 1970. Given the observed education distributions, the estimated regression coefficients, and the first stage estimates of the residual wage distribution, it is possible to uncover the probability density of total wages for all three racial groups in 1970. Let this be given by $\widehat{x_{70}^{j}}(w)$ for $j = w, b, h$.

The parameters m_1 and m_2 are estimated by imposing the equilibrium requirement that the probability that parents invest be equal to firms' prior beliefs about investment. That is, equation (4) must hold in equilibrium. By substituting the parametric specification of $G(c|w)$ into (4) for blacks, whites and Hispanics, it is possible to solve for m_1 and m_2. In particular, the following moment conditions are used to estimate m_1 and m_2 by least squares.

$$\widehat{\pi_{90}^{j}} = m_1 \int_w \frac{1}{w} \widehat{x_{70}^{j}}(w) dw + m_2 B(\widehat{\pi_{90}^{j}}), j = w, b, h.$$

6 Data

The data are from the 1970 and the 1990 U.S. Census. In both years, the data are from a 1-in-100 national random sample of the U.S. population. These data are available through the University of Minnesota's Integrated Public Use Microdata Series (Ruggles, 1997).

Both the 1970 and the 1990 sample are restricted in a number of ways. First, they include only men since women often have a weak labor force attachment. In addition, individuals must have worked at least one week during the year and have a positive wage and salary income in order to be included.

A number of restrictions are also made so that the individuals considered in 1970 are likely to represent the parents of individuals considered in 1990. First, in 1990 only individuals between the ages of 35 and 45 are included in the sample. Then, assuming that a generation is 25 years long, the parents of those workers would have been between the ages of 40 and 50 in 1970, and so the 1970 sample is restricted accordingly. These age restrictions also have the advantage of capturing workers near the top of their age-earnings profile. In addition, both the 1970 and the 1990 sample are restricted to include only those individuals who were born in the United States. This also avoids possible biases introduced by including non-native English speakers.

Finally, individuals are classified into one of three racial groups: whites, blacks and Hispanics. In particular, individuals are classified as white if they identify themselves as both white and non-Hispanic. Individuals are classified as black if they identify themselves as black and either Hispanic or non-Hispanic. And, lastly, Hispanics are those who identify themselves as both white and Hispanic. All other racial groups are excluded from the sample.

Total wages are calculated by dividing annual wage and salary income by the number of weeks worked that year. I do not use hourly wage data because hourly wage data is not available in 1970. All dollar values have been deflated to the 1989 level.

Constructing Quantiles of the Residual Wage Distribution

In both years, for all three racial groups, I calculate the 81 percentiles starting at the 10th percentile and ending at the 90th percentile. I do not use quantiles below the 10th percentile or those above the 90th percentile so that the results are not driven by outliers in the data. In both years individuals with top coded earnings have residual wages that lie above the 90th percentile of the residual wage distribution for that individual's group. Thus, in effect, their wages are not used in estimation.

7 Results

Stage 1 Results

The results of the first stage of estimation are displayed in Table 1. As might be expected, individuals who finish college earn higher wages than those who do not, and this benefit is higher for whites than for blacks and Hispanics. In addition, for all three groups, the return to college declines from 1970 to 1990. These estimates are used to construct the residual wage distribution for each racial group.

Table 1: Stage 1 Estimates

Dependent Variable=Weekly Wage

(Standard Errors in Parentheses)

Year=1970	Whites	Blacks	Hispanics
College Wage Premium	404.44	356.71	347.27
	(3.71)	(15.87)	(26.92)
Age	28.42	54.76	-3.95
	(14.26)	(34.14)	(58.03)
Age Squared	-.299	-.616	.020
	(.158)	(.380)	(.647)
Mean Sq. Dev.	398.59	293.19	278.23
N	81,487	7,407	2,381

Year=1990	Whites	Blacks	Hispanics
College Wage Premium	286.02	226.36	256.43
	(4.78)	(15.76)	(15.11)
Age	3.57	.247	6.08
	(20.53)	(51.36)	(50.08)
Age Squared	.089	.111	-.001
	(.258)	(.645)	(.630)
Mean Sq. Dev.	795.47	668.14	383.98
N	126,025	13,886	4,837

Stage 2 Results

The results of the second stage of estimation and the bootstrapped standard errors are displayed in Table 2. In both 1970 and 1990, the probability that blacks and Hispanics are qualified is lower than the probability that whites are qualified. This result follows necessarily from the set up of the model since the model interprets racial wage differentials as differences in unobservable investment.

Table 2: Quantile Matching Estimates

	1970			1990		
	White	Black	Hispanic	White	Black	Hispanic
π	.3703	.2240	.2967	.3111	.2366	.2596
	(.0042)	(.0037)	(.0052)	(.0024)	(.0026)	(.0034)
γ_q	1.478	1.835	1.680	1.580	1.708	1.675
	(.0070)	(.0229)	(.0330)	(.0065)	(.0152)	(.0266)
γ_u	1.612	1.290	1.356	1.610	1.447	1.476
	(.0104)	(.0138)	(.0275)	(.0081)	(.0128)	(.0301)
y_u		82.15			49.79	
		(1.86)			(1.51)	
y_q		1,543.14			1,804.41	
		(17.18)			(14.55)	
Root MSE		10.31			7.48	

The numbers in Table 2 indicate a substantial increase in the returns to skill between 1970 and 1990. Over that 20 year period, the difference between the productivity of qualified workers and unqualified workers, $y_q - y_u$, grew 20% from \$1,460.99 to \$1,754.62. This increase reflects the fact that there has been an increase in the dispersion of wages, controlling for college completion.

Interestingly, in the face of these increasing returns to skill, the proportion of qualified workers is found to decline from 1970 to 1990. This result reflects the fact that while there was growth at the top end of the wage distribution between 1970 and 1990, the median wage stayed relatively constant. The model interprets this shift as an increase in the productivity of qualified workers combined with a decrease in the proportion of qualified workers. Given that there have been large increases in observable investment (e.g. years of schooling) from 1970 to 1990, the finding that unobservable skill has declined may seem strange. However, the fact that observable investment has increased does not necessarily imply that unobservable investment should have also increased.

In addition to the racial differences in investment, there are also substantial differences in the parameters governing the distributions of the productivity signals across the three racial groups. In particular, in both 1970 and 1990 $\gamma_q > \gamma_u$ for blacks and Hispanics, but $\gamma_q < \gamma_u$ for whites. The implication is that employers receive relatively imprecise signals from unqualified black and Hispanic workers and relatively precise signals from unqualified white workers. As it turns out, this tends to disadvantage blacks and Hispanics because, as the Stage 2 estimates indicate, firms believe that most workers are unqualified. Thus, since unqualified blacks and Hispanics have relatively noisy signals, firms are unlikely to be "impressed" when workers from these groups have high signals, and, as a result, it is difficult for qualified blacks and Hispanics to distinguish themselves.

Since racial differences in the signaling distributions will turn out to explain a large portion of racial wage differentials it is worth examining what features of the wage distribution γ_q and γ_u capture. In general, γ_q and γ_u capture both the variance and the skewness of the wage distribution. In particular, the higher γ_q and γ_u, the higher the variance of the wage distribution. In addition, as $\frac{\gamma_q}{\gamma_u}$ increases, the wage distribution becomes more positively skewed. The reason is that as γ_q increases and γ_u decreases, the distribution of productivity signals for qualified workers becomes increasingly concentrated near one, and the distribution of productivity signals for unqualified workers becomes increasingly dispersed. Thus, since wages are a monotonic function of a worker's signal, the wage distribution becomes more positively skewed.

Apart from γ_q and γ_u, the only other parameter that affects skewness is π and, as π increases, the skewness of the wage distribution also increases. However, as mentioned above, π also reflects the mean of the wage distribution. Thus, the fact that blacks have mean lower wages than both whites and Hispanics in both 1970 and 1990, is reflected in the fact that π^b is less than π^w and π^h in both years. Moreover, given these racial differences in π, the wage distribution for blacks and Hispanics should be more negatively skewed than the wage distribution for whites. However, as it turns out, the difference in the skewness of these distributions is relatively small, and the model interprets this as being the result of differences in the shape of the underlying signaling distribution, and in particular, that $\frac{\gamma_q}{\gamma_u}$ is larger for blacks and Hispanics than it is for whites.

In order to examine how well these parameter estimates in Table 3 replicate the observed wage distribution, 1,000 wages are simulated according to the estimated signal distributions and wage schedules. Table 3 then compares summary statistics from the estimated and observed wage distributions. In addition, the histograms at the end of the paper plot both the predicted and the actual wage distributions for each of the three racial groups side by side for comparison. As can be seen, the parameter estimates do a reasonably good job of capturing the shape of the wage distribution.

Table 3a: Goodness of Fit, 1970 Residual Wage Model

	Whites		Blacks		Hispanics	
	Actual	Estimated	Observed	Estimated	Observed	Estimated
Mean	617.00	623.70	425.75	410.75	523.95	517.83
St Dev.	259.76	265.42	202.87	193.51	232.73	225.32
Skewness	.683	.681	1.051	.546	.734	.441
Kurtosis	3.64	3.04	5.19	3.01	4.01	2.94

Table 3b: Goodness of Fit, 1990 Residual Wage Model

	Whites		Blacks		Hispanics	
	Actual	Estimated	Observed	Estimated	Observed	Estimated
Mean	509.07	579.61	471.17	469.98	509.74	516.12
St Dev.	310.24	306.40	266.44	268.32	278.32	278.57
Skewness	0.980	0.755	1.199	.855	1.091	.662
Kurtosis	4.27	3.26	5.32	3.56	4.88	3.10

Stage 3 Results

Tables 4 and 5 present the results from the third stage of estimation. The fact that $m1 > 0$ confirms the negative relationship between parental investment costs and wages.

Table 4: Parameters of the Cost Distribution

(Standard Errors in Parenthesis)

m_1	5.549 (2.73)
m_2	0.0012 (.00002)

As discussed above, costs are uniformly distributed on $[\frac{m_1}{m_2 w}, \frac{w+m_1}{m_2 w}]$. These estimates indicate that investment costs are uniformly distributed over an interval that is over $833 long. So, for example, the benefit of investing would have to increase by approximately $83 per week for the probability that parents invest to increase by 10 percent.

Table 5: Implied Costs and Benefits

	White	Black	Hispanic
$E[c\|invest]$	130.87	100.77	103.69
$B(\pi)$	270.14	216.04	218.56
PDV $E[c\|invest]$	41,334	31,827	32,749
PDV $B(\pi)$	85,321	68,234	69,030

Using the Stage 3 estimates, Table 5 presents the average investment cost for white parents and black parents conditional on the fact that they invest. These costs are weighed against the benefit of investment. Up until now, that benefit has been expressed in terms of increases in a worker's expected weekly wage. However, parents really care about the increase in their children's expected lifetime earnings, and the cost of investment should also be considered on this scale. Thus, Table 5 displays the present discounted value of the benefit of investment, and accordingly re-scales the average cost of investment for parents who invest. It is assumed that workers work 50 weeks a year for 40 years and that the annual interest rate is 3%. Note that conditional on investing, the cost of investment is lower for blacks and Hispanics than it is for whites. This is because the benefit of investment is lower for blacks and Hispanics than it is whites.

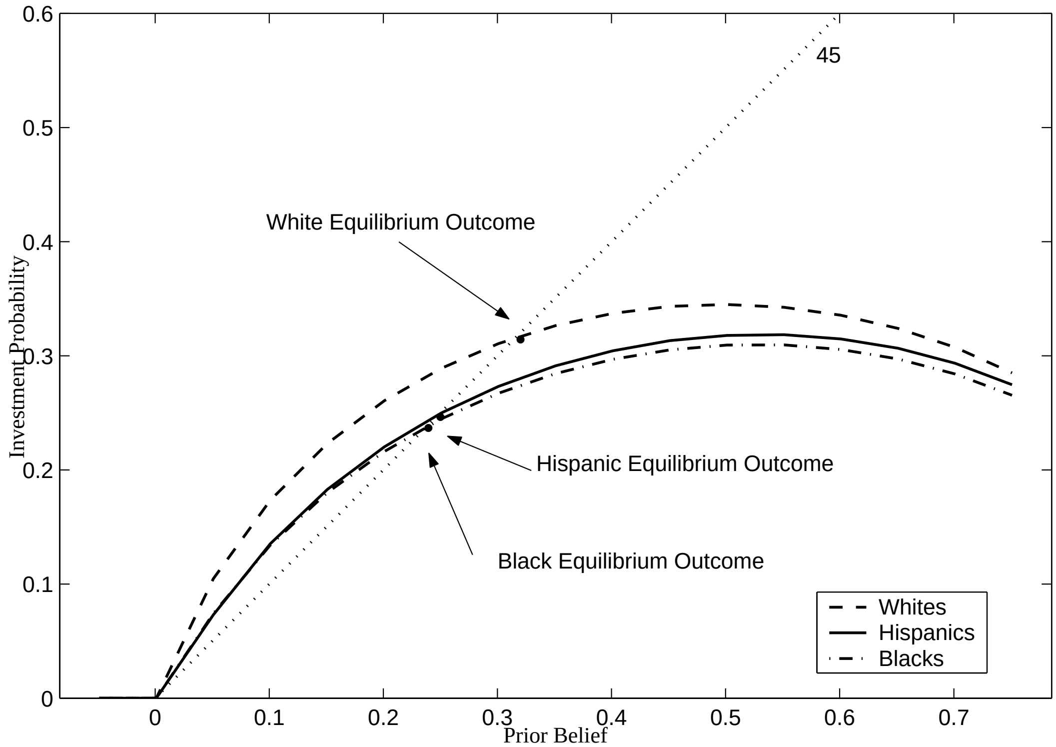


Figure 2: All Possible Equilibria in 1990

The Sources of Racial Wage Inequality

In order to understand the sources of racial wage inequality, the parameter estimates are first used to calculate all of the model's equilibria, including those that are not realized. By doing so, it is possible to determine whether coordination failures were responsible for racial wage disparities in 1990. Figure 2 depicts all of the model's equilibria in 1990. The arced lines represent the probability that parents who are playing a best response invest as a function of the firms' prior beliefs. Since Hispanics earn lower wages than whites, and since blacks earn lower wages than Hispanics, the probability that Hispanics invest is lower than the probability that whites invest, and the probability that blacks invest is lower than the probability that Hispanics invest at every level of π. Since the firms' equilibrium beliefs are self-confirming, the intersections of the parents' best response functions with the 45-degree line represent all of the equilibria of the model.

According to the parameter estimates, the equilibrium outcome for all three racial groups is such that there exists no other equilibrium in which any group could have had a higher investment level. In addition to the realized equilibria, the only other equilibria are clustered around zero. Thus, the results indicate that racial wage inequality in 1990 is not the result of a coordination failure.

In order to understand what features of the data lead to the conclusion that coordination failures are not responsible for racial inequality, consider the following example. Suppose that there are two groups who are identical in all ways except that they have different equilibrium investment levels. According to the model, that difference in investment could only be explained by a coordination failure. However, in this model, blacks, whites and Hispanics may differ along a number of dimensions. In particular, there may be racial differences in the ability of workers to accurately signal their productivity ($\gamma_q^w \neq \gamma_q^b \neq \gamma_q^h$ and $\gamma_u^w \neq \gamma_u^b \neq \gamma_u^h$), and there may be differences in the distribution of wages in the parents' generation. These factors alone may lead to different investment levels across the three groups.

In this context, in order to find evidence of a coordination failure, it would have to be the case that racial differences in the precision of the productivity signals and in the distribution of the parental wages could not account for a substantial portion of racial differences in investment. In order to examine whether this is the case, Column 3 of Table 6 reports the simulated probability that parents from each group would have invested if all three groups had the same productivity signal precision as whites in 1990, and Column 4 reports the simulated probability that parents from each group would have invested if the 1970 wage distribution had been identical to the white wage distribution for all three groups. Finally, Column 5 reports the combined effect of equalizing both the precision of the signals and the wage distribution across all three racial groups. As a point of comparison, Column 2 reports the estimated investment probabilities from the second stage of estimation. As can be seen, once racial differences in parental wages and in the distribution of productivity signals are accounted for, all three racial groups invest at nearly the exact same rate. The remaining discrepancy can be explained by differences in the benefit to investing that result from differences in the firms' prior beliefs. However, because this remaining difference is small, the relationship between differences in investment levels and differences in the firm's prior beliefs is found to be weak, and the model interprets this as evidence against the presence of a coordination failure.

In terms of the data, this implies that racial differences in the mean wage are not large relative to racial differences in parental wages and in the parameters of the distribution of productivity signals (which reflect racial differences in the variance and skew of the wage distribution). Had the difference in the mean of the wage distributions for the three groups been substantially larger, then the model would have interpreted this as the result of a coordination failure.

Table 6: Simulated Investment with White Signaling Ability and with White 1970 Wage Distribution

	Stage 2 Estimated Investment	White Signaling Precision	White 1970 Wage Distribution	White Signaling Precision and 1970 Wages
Black	.2366	.2747	.2499	.2825
Hispanic	.2596	.2903	.2629	.2941
White	.3111	.3111	.3111	.3111

Table 6 also makes it clear that differences in the precision of the productivity signals play the most significant role in explaining racial differences in parental investment. Whereas differences in signaling precision account for nearly 52 percent of the black-white investment differential, differences in wages explain only 17.9 percent. Similarly, for Hispanics, nearly 60 percent of the Hispanic-white investment differential is explained by differences in signaling precision between the two groups, but only 6.4 percent of the difference is explained by wages. Moreover, the fact that similar informational asymmetries are found in 1970 suggests that signaling precision may play an important role in explaining persistent racial wage differentials.

As mentioned earlier, racial differnces in signaling precision lead to lower investment levels for blacks and Hispanics because unqualified workers from these groups have relatively noisy productivity signals. Thus, since firms believe that most workers are unqualified, it is difficult for qualified blacks and Hispanics to distinguish themselves from unqualified workers who happen to have high signals. This tends to lower the incentives for blacks and Hispanics to invest and leads to lower equilibrium investment levels.

It should be noted that the intergenerational correlations in earnings implied by the parameter estimates is substantially below conventional estimates. For example, the implied intergenerational correlation in earnings for whites is less than .1 whereas both Solon (1992) and Zimmerman (1992) estimate that this correlation is closer to .4. Thus, the model appears to do a poor job in capturing the transmission of earnings across generations, and the importance of past inequality in explaining current racial wage differentials may be substantially understated by this model. The source of this inconsistency may be due to the strong functional form assumptions made about the distribution of parental investment costs conditional on parental wages, and although attempts were made to estimate more flexible functional forms using linked parent-child data from the Panel Survey of Income Dynamics (the PSID), the sample sizes in that data set were too small to generate meaningful estimates.

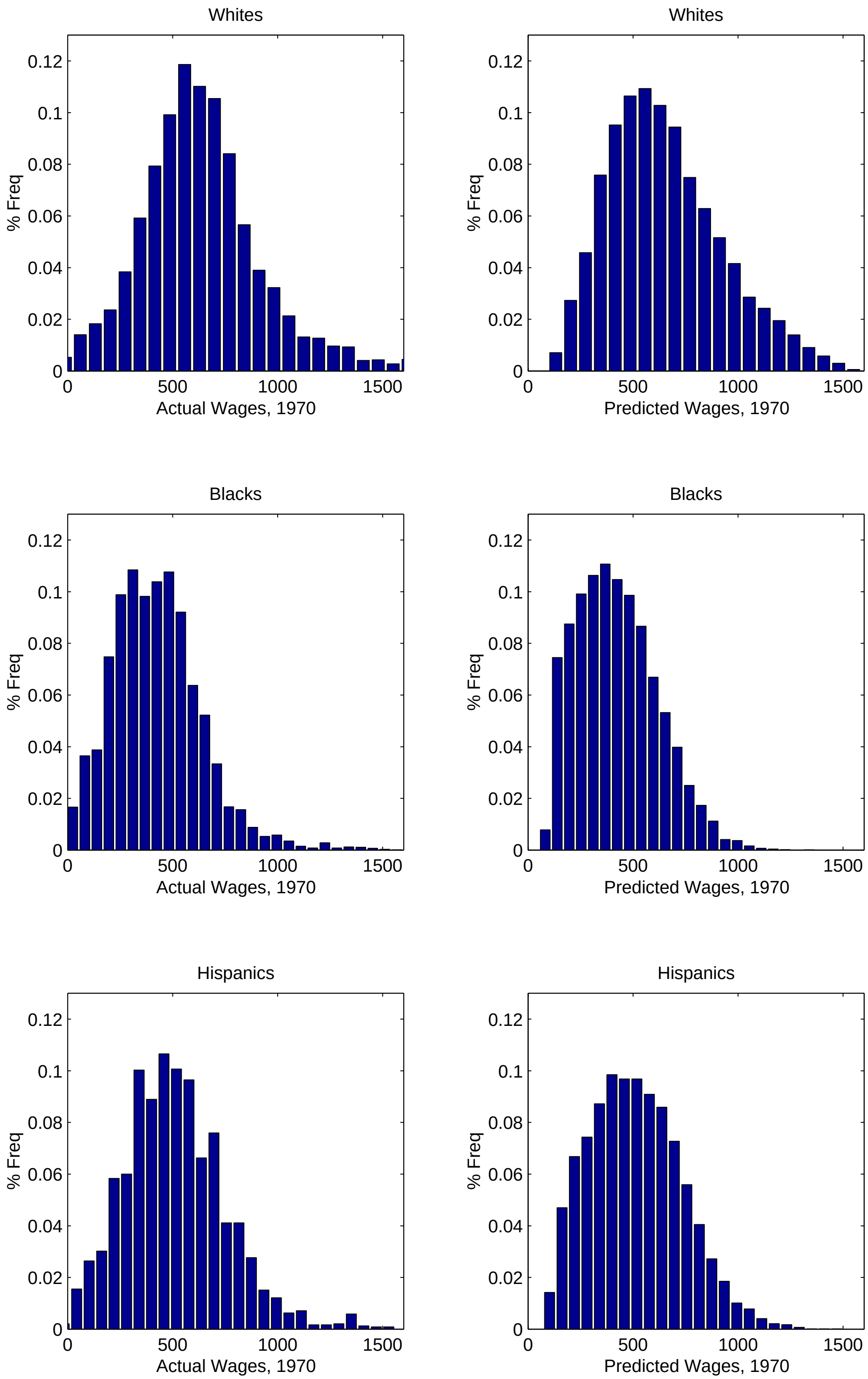
8 Conclusion

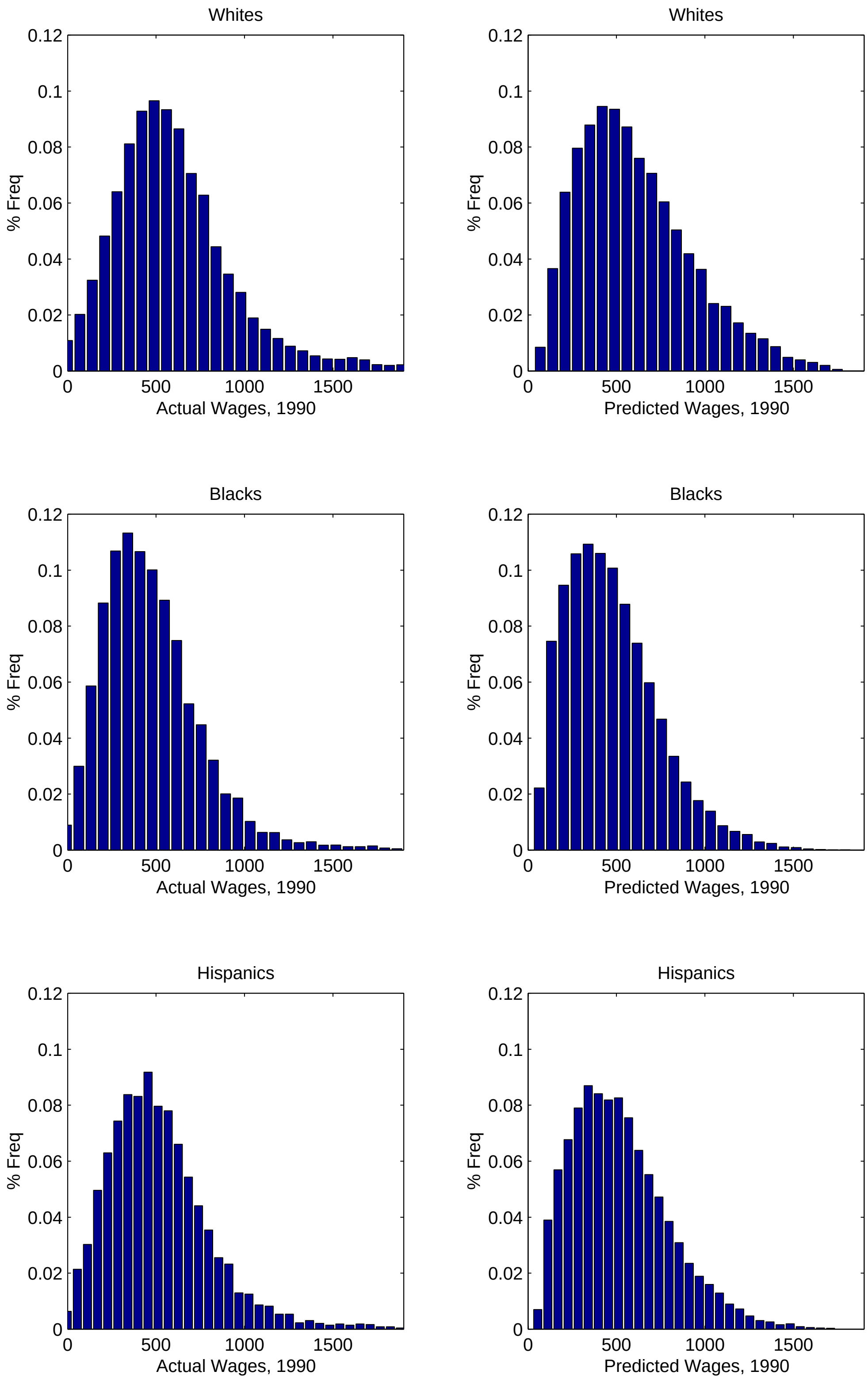
In the model developed in this paper, incomplete information is the root cause of racial discrimination. In this context, disparate outcomes can arise due to the existence of multiple equilibria, past income inequality or racial differences in the precision with which workers are able to signal their productivity. This paper then examines which of these three phenomena is primarily responsible for racial wage inequality by estimating the model's fundamental parameters with data from the 1970 and 1990 U.S. Census.

According to the estimates, the model indicates that racial inequality in 1990 was primarily a result of racial differences in the precision with which workers from different racial groups signal their productivity. This informational asymmetry reduces the benefit to investing for blacks and Hispanics, and, as a result, black and Hispanic parents invest less than white parents, and firms rationally expect black and Hispanic workers to be less productive than white workers. Moreover, the fact that similar informational asymmetries are found to exist in 1970 suggests that racial differences in a group's ability to reliably signal their productivity may play an important role in explaining racial wage inequality in 1970 as well. Since this is one of the only papers to directly estimate racial differences in signaling precision and to link those differences with differential investment incentives, this research provides some of the first direct evidence that informational asymmetries may play an important role in explaining the persistence of black-white wage differentials. Given this finding, additional work that attempts to examine the magnitude of these informational asymmetries and their impact on wages and the return to human capital seems warranted.

The estimates also indicate that past income inequality does not play an important role in accounting for racial wage differentials. However, it should be noted that the correlation in earnings across generations implied by the parameter estimates is well below conventional estimates, and, thus, there is reason to suspect that the model understates the importance of past inequality in explaining future disparities in earnings.

Finally, as in Moro (2001), the empirical estimates do not support the hypothesis that racial wage inequality is the result of multiple equilibria. However, discerning the presence of multiple equilibria may be difficult, and some of the assumptions needed to gain identification may be obscuring their presence. First, the fact that investment is a binary choice allows the distribution of productivity signals to be easily identified. Ideally, a less restrictive approach would estimate a worker's productivity directly. Then, under the assumption that the discrepancy between the worker's wage and their estimated productivity was the result of signal noise, the distribution of productivity signals conditional on a worker's productivity could be estimated non-parametrically. The difficulty, of course, is in finding a reliable measure of worker productivity. Second, functional form assumptions are also used to identify the distribution of parental investment costs, and although attempts were made to estimate these distributions nonparametrically, these efforts were not successful due to the lack of data sets with a large number of repeated observations on linked parent-child pairs. Thus, the multiple equilibria story may not be dead. We simply may need richer models or better data to pick up its effects.


Whites
Whites
% Freq
0.12
0.1
0.08
0.06
0.04
0.02
0
0
500
1000
1500
Actual Wages, 1970
Predicted Wages, 1970
Blacks
Blacks
Actual Wages, 1970
Predicted Wages, 1970
Hispanics
Hispanics
Actual Wages, 1970
Predicted Wages, 1970


Whites
% Freq
Actual Wages, 1990
Whites
% Freq
Predicted Wages, 1990
Blacks
% Freq
Actual Wages, 1990
Blacks
% Freq
Predicted Wages, 1990
Hispanics
% Freq
Actual Wages, 1990
Hispanics
% Freq
Predicted Wages, 1990
0
0.02
0.04
0.06
0.08
0.1
0.12
0
500
1000
1500

Appendix A: Proofs

Proof of Proposition 1: Begin by proving the following intermediate lemma.

Lemma 1 *Suppose $\{\mathbf{w}_{t,i}\}_{i=1,2}$ is a pair of best responses at time t, then for each group j at every time t, $w^j_{t,1}(\theta) = w^j_{t,2}(\theta)$ for almost all $\theta \in [0, 1]$.*

Proof: Suppose that for some group j at time t there is a positive measure set $\Theta^j_t \subset [0, 1]$ on which firms post different wages so that for some i, $w^j_{t,i}(\theta) > w^j_{t,i'}(\theta)$ for all $\theta \in \Theta^j_t$. Then by reducing wages on this set while keeping them above the wages posted by the other firm, this firm can increase profits since output will not change, but the total wage bill will be smaller.

Proof of Proposition 1 (continued): Lemma 1 establishes that for each j, $w^j_{t,1}(\theta) = w^j_{t,2}(\theta) = w(\theta, pi^j_t)$ Suppose that $w^j_t(\theta) < w^j_t(\theta, \pi^j_t)$ for a positive measure set $\tilde{\Theta}^j \subset [0, 1]$. Consider an alternative strategy $\hat{w}_{t,1}$ for firm 1 where $\hat{w}^{j'}_{t,1}(\theta) = w^{j'}_t(\theta)$, and $\hat{w}^j_{t,1}(\theta) = w^j_t(\theta) + \epsilon$ for $\theta \in \tilde{\Theta}^j$ for some $\epsilon > 0$ while $\hat{w}^j_{t,1}(\theta) = w^j_t(\theta)$ for $\theta \in [0, 1] \setminus \tilde{\Theta}^j$. All workers from group j at time t whose test signal $\theta \in \tilde{\Theta}^j$ will then accept firm 1's offer. The difference in profits between $\{\hat{\mathbf{w}}_{t,1}\}$ and $\{\mathbf{w}_{t,1}\}$ is:

$$\int_{\theta \in \tilde{\Theta}^j} \left\{ \frac{1}{2} \left[w^j_t(\theta, \pi^j_t) - w^j_t(\theta) \right] - \epsilon \right\} f_{\pi^j_t}(\theta) d\theta,$$

which is strictly positive if ϵ is sufficiently small. Thus, if $w^j_t(\theta) < w^j_t(\theta, \pi^j_t)$ player 1 has a profitable deviation. This contradicts Lemma 1. A parallel argument can establish the fact that a profitable deviation exists if $w^j_t(\theta) > w^j_t(\theta, \pi^j_t)$.

Proof of Proposition 2: If $f_q(\cdot)$ and $f_u(\cdot)$ are continuous on $[0, 1]$, then $w(\cdot, \cdot)$ is continuous on $[0, 1] \times [0, 1]$ which, in turn, implies that $B(\cdot)$ is continuous on $[0, 1]$. Thus, since $g(\cdot|w)$ is assumed to be continuous on $[\underline{c}, \overline{c}]$ for all w, we know that $G^j_{t-1}(B(\cdot))$ is continuous on $[0, 1]$. By the continuity of $B(\cdot)$ and $G^j_{t-1}(\cdot)$, the fact that $B(0) = 0$ and that fact that investment costs are strictly positive, we know that there exists some $\epsilon > 0$ such that $G^j_{t-1}(B(\epsilon)) = 0$. Thus, since $G^j_{t-1}(B(\eta)) > \eta$ and since $G^j_{t-1}(B(\cdot))$ is continuous, we know by the Intermediate Value Theorem, that there exists some $\pi' \in (0, \eta)$ such that $G^j_{t-1}(B(\pi')) = \pi'$. Similarly, since $B(1) = 0$, then by the same logic as above, we know that there also exists a solution on $(\eta, 1)$

References

[1] ALTONJI, J. and PIERRET, C. (2001), "Employer Learning and Statistical Discrimination", *Quaterly Journal of Economics*, **116**, 313-350.

[2] ARROW, K. (1972), "Models of Job Discrimination" and "Some Mathematical Models of Race in the Labor Market", in A. Pascal (ed.) *Racial Discrimination in Economic Life* (Lexington, MA: Lexington Books).

[3] ARROW, K. (1973), "The Theory of Discrimination", in Orley Ashenfelter and Albert Rees (eds.), *Discrimination in Labor Markets* (Princeton, NJ: Princeton University Press).

[4] BECKER, G. and TOMES, N. (1986), "Human Capital and the Rise and Fall of Families", *Journal of Labor Economics*, **4**, S1-S47.

[5] CARD, D. and LEMIEUX, T. (1996), "Wage Dispersion, Returns to Skill, and Black-White Wage Differentials", *Journal of Econometrics*, **74**, 319-361.

[6] CHAMBERLAIN, G. (1994), "Quantile Regression, Censoring, and the Structure of Wages", in C. Sims (ed.), *Advances in Econometrics: Sixth World Congress, Volume 1* (New York, NY: Cambridge University Press).

[7] COATE, S. and LOURY G. (1993), "Will Affirmative Action Eliminate Negative Stereotypes?", *The American Economic Review*, **83**, 1220-1240.

[8] CRAMER, H. (1951), *Mathematical Methods of Statistics* (Princeton, NJ: Princeton University Press).

[9] DONOHUE, J. J. III and HECKMAN, J. (1991), "Continuous Versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks", *Journal of Economic Literature*, **29**, 1603-1643.

[10] EPPLE, D. and SIEG, H. (1999), "Estimating Equilibrium Models of Local Jurisdictions", *Journal of Political Economy*, **107**, 645-681.

[11] LOURY G. (1981), "Intergenerational Transfers and the Distribution of Earnings", *Econometrica*, **49**, 843-867.

[12] LUNDBERG, S. and STARTZ, R. (1983), "Private Discrimination and Social Intervention in a Competitive Labor Market", *The American Economic Review*, **73**, 340-347.

[13] MCCALL, J. (1972), "The Simple Mathematics of Information, Job Search and Prejudice", in A. Pascal (ed.), *Racial Discrimination in Economic Life* (Lexington, MA: Lexington Books).

[14] MORO, A. (2001), "The Effect of Statistical Discrimination on Black-White Wage Differentials: Estimating a Model with Multiple Equilibria", International Economic Review, forthcoming.

[15] PHELPS, E. (1972), "The Statistical Theory of Racism and Sexism", *American Economic Review*, **62**, 659-61.

[16] RUGGLES, S. and SOBEK, M. et. al. (1997), *Integrated Public Use Microdata Series: Version 2.0 [On-line]*, (Minneapolis: Historical Census Projects, University of Minnesota), URL: http://www.ipums.umn.edu.

[17] SOLON, G. (1992), "Intergenerational Income Mobility in the United States", *American Economic Review*, **82**, 393-408.

[18] ZIMMERMAN, D. (1992), "Regression Toward Mediocrity in Economic Stature", *American Economic Review*, **82**, 409-429.