CARESS Working Paper #95-06
Statistical Discrimination, A±rmative Action, and Mismatch

Jaewoo Ryoo[¤]
Department of Economics
University of Pennsylvania

October 1994; Revised March 1995

Abstract

This paper analyzes the economic consequences of a±rmative action in the presence of statistical discrimination. In the model, workers with di®ering abilities have comparative advantages in jobs with di®ering complexities. Employers, having a biased belief on the ability of minority workers, require higher credentials when promoting them to more productive jobs, which discourages their human capital investment. When a±rmative action policy is enforced, some under-quali¯ed minority workers are promoted to di±cult jobs. Those workers, as well as some majority workers who are over-quali¯ed for, but have to take, easy jobs lose because their comparative advantages are not utilized. This ine±ciency due to mismatch is not necessarily outweighed by the long term gain brought about by the policy, if groups di®er substantially in their human capital investment costs. Appropriately reinterpreted, the model explains why drop-out rates and the returns to college education di®er between blacks who attend black and non-black colleges.

[¤]Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104. Phone; (215) 898-7711. E-mail; jryoo@ssc.upenn.edu. I thank Stephen Coate, Adrian Masters, Stephen Morris, Sherwin Rosen, Paul Taubman, and the participants of combined workshop of Development, IO, Labor, and Political Economy at Penn for helpful comments.

1. Introduction

Overview

This paper builds a model of statistical discrimination in an overlapping generations framework, and assesses the economic (in)e±ciency of a±rmative action. In the model every worker is assigned to the same job in the ¯rst period, and produces a \signal" that is noisy but increasing in the \innate" ability[1] and the investment/e®ort of the worker. Based upon the signal, in the second period workers are assigned to two di®erent jobs within a ¯rms's hierarchy: a more di±cult but more productive higher level job and an easier but less productive lower level job. A worker with a higher innate ability and a larger human capital, henceforth called an \abler" worker, is more productive in any job. The less able, however, have a comparative advantage in an easier - less productive job. A worker is uncertain regarding his own ability and chooses a job which gives a higher expected payo®, based upon his believed level of ability. Employers have a downward biased belief about the ability of minority workers, and require higher credentials in order for them to qualify for more productive jobs. This discourages productivity enhancing human capital investment by the disadvantaged workers. As a consequence, the ex post distribution of the ability of minority workers is lower than that of the majority workers, even though their ex ante distribution of innate ability is the same.

A±rmative action policies considered in the current context enforce an equal representation of workers across groups in a more productive job. Those policies can have an unwanted side e®ect, i.e., production ine±ciency, because they induce a mismatch between the tasks of di®ering complexities and the workers of di®ering abilities. When a minority worker is promoted, he does not know if he was promoted with or without the help of a±rmative action. If the worker is under-quali¯ed but is o®ered a di±cult job he may

[1] \Innate" ability refers to the productivity enhancing skills which are not acquired through human capital investment. For example, the ability measured by IQ may be in°uenced by prenatal care, intellectual stimuli given to infants, as well as by genes. As long as that kind of ability is not determined by the choice of workers, it is conveniently included within innate ability.

become unduly optimistic about his own ability and accept the o®er. Consequently, the worker's performance is lower than it would be in a more suitable job. This misallocation of workers is essentially due to the fact that a±rmative action confuses signals. Of course, the mismatch e®ect also works in the other direction as \reverse discrimination"; An over-quali¯ed majority worker is forced to take an easy job that underutilizes his productive ability. The real question then is whether a±rmative action also creates bene¯ts which outweigh matching ine±ciency. The answer to this question is shown to depend much upon the investment cost deferential across groups.

Even though the theory has a broad applicability to the issues involving the allocation of workers in labor markets in general, it is particularly pertinent to explain why the drop-out rates are lower and the returns to college education higher for the blacks who attend historically black colleges (HBCs) when compared to the blacks who attend non-black colleges. Reinterpreting the ¯rst period as the pre-college period when students produce \test scores" upon which college admissions are based, and the di±cult and easy jobs as the \prestigious" and \less prestigious" colleges, prestigious colleges augment earning power more, but only for those who have higher levels of ability. The less able students may fare better at less prestigious colleges. When a±rmative action stipulates that the prestigious colleges have a student population proportional to the population as a whole, some of the black students who have acquired less human capital (as well as lower innate ability) will be admitted. The model implies that those students will perform more poorly and have an increased chance of dropping out. Also, given that the performance in college is as important as the prestigiousness of the school attended in determining earnings, the returns from college education may be lower for blacks who attend non-HBCs.

Related Literature

The paper deals with the topics of statistical discrimination and a±rmative action. Regarding statistical discrimination, there are two lines of research with slightly di®erent °avors. One (e.g., Arrow (1973), Coate and Loury (1993b)) deals with the case when employers have a biased belief about the productivity of minority workers. Another one (e.g., Phelps (1972), Borjas and Goldberg (1979), Lundberg and Startz (1983), Milgrom and Oster (1987), and Lundberg (1991)) deals with the case where employers can assess the productivity of minority workers less precisely. The current model is general enough to accommodate Phelps-type discrimination as well as, but more concerned with, Arrow-type discrimination. Yet, there are important differences between the current model and the models of Arrow (1973) and Coate and Loury (1993 b) (henceforth called ACL).

A crucial assumption in ACL is that employers have a biased belief about the fraction of minority workers who are qualified for high-paying skilled jobs. A worker who has invested and is qualified for a skilled job may not be assigned to it because of the noisy signal and the biased belief of employers. Note that the existence of discriminatory equilibria in ACL is not robust to alternative wage contracts. For example, a wage contract where a worker posts a bond that is forfeited to the firm if the worker turns out to be unqualified can eliminate discriminatory equilibria and result in an efficient outcome. This suggests that the employment contract in ACL may not be supported by the market. The current model, in contrast, treats the statistical discrimination problem in a more standard (and appealing) way: employers have a biased belief about minority workers' abilities.[2] Since workers are assumed to be uncertain about their own abilities, there is no room for a contractual improvement.

More importantly, the economic consequences of affirmative action in the presence of statistical discrimination are analyzed in the current model in an environment where workers with differing abilities have comparative advantages in different jobs. This enables a direct assessment of the economic gain/loss brought on by affirmative action.

Not much effort has been made to assess the economic consequences in spite of the importance of affirmative action as an institution governing the labor markets. Among a few exceptions is Welch (1976), but the skill distribution of workers is exogenous in his

[2] There exists a wide gulf between private belief and public discussion on the issues involving the abilities of minority workers, and few employers would publicly admit a negative belief about the ability of minority workers. Still this assumption might describe the current state well. The confusion among the public and the big media debate provoked by a new book The Bell Curve: Intelligence and Class Structure in American Life by Charles Murray and Richard Herrnstein is a proof of the sensitive and controversial nature of those issues.

model, (statistical) discrimination is not explicitly modeled, and the welfare loss due to employment quotas for minorities is not weighed against the loss due to discrimination. Lundberg (1991) treats mismatch effects resulting from affirmative action, but without a well-grounded theoretical justification. Schotter and Weigelt (1992) report that affirmative action programs always benefit disadvantaged groups. This paper shows that is not always the case. It is the first attempt to determine whether affirmative action results in an efficient outcome by incorporating statistical discrimination, workers' human capital investment incentives, comparative advantages of workers in jobs of differing complexities, and affirmative action in a unified framework. The framework enables a direct comparison of the welfare loss due to discrimination in the absence of affirmative action policy and the welfare loss due to mismatch in the presence of the policy.[3]

The basic model presented in the next section describes the firms' optimal job assignment policy and the workers' optimal investment rule. The market equilibrium is characterized in Section 3. Section 4 analyzes the economic consequences of statistical discrimination and affirmative action. A conclusion follows in Section 5.

2. The Basic Model

Two easily identifiable groups, B and W, indicating minority and majority groups respectively, exist in the economy, and each group has a continuum of risk neutral workers. Workers live for two periods only, and worker generations are identical and overlap. There are many risk neutral firms which live infinitely. Each firm hires a small fraction of workers. A worker's expected life-time utility is given by,

$$E\left[\sum_{t=1}^{2} \beta^{t-1}(w_t - C(e_t))\right]; \qquad (2.1)$$

where E is an expectation operator, β is the discount rate, and w_t is the wage at period t. $C(e_t)$ is the cost of effort (e) at t, with $C(0) = 0$, and $C'(e) > 0$ and $C''(e) > 0$ for any $e \geq 0$.

[3] The mismatch effects addressed in this paper suggest that previous research, by ignoring such effects, may have exaggerated the economic gains of blacks brought on by affirmative action. See also Welch (1976).

In $t = 1$, everybody works at the same job. Worker i with "innate" ability a_i and non-negative investment/effort e_{1i} produces the first period output (q_{1i}) according to

$$q_{1i} = a_i + e_{1i} + \epsilon_{1i}; \tag{2.2}$$

where $\epsilon_{1i} \sim N(0; \sigma^2_{\epsilon_1})$ is the noise. The first period output may or may not be observed by workers: On the other hand, firms have a conjecture on e_1; observe q_1; and estimate the innate ability of a worker using the two. Based on that estimate, firms assign workers to either difficult (D) or easy (E) jobs in the second period. Hence q_1 serves as a "signal" upon which assignment of workers to different jobs is based.

The second period output at job j, $j = D; E$ is

$$q_{ji} = \alpha_j + \gamma_j(a_i + e_{1i} + e_{ji}) + \epsilon_{ji}; \tag{2.3}$$

where e_j is the effort in job j, and ϵ_j is the noise in the second period output in job j. Assume that $\epsilon_{ji} \sim N(0; \sigma^2_{\epsilon_j})$ and $Cov[\epsilon_1; \epsilon_j] = Cov[\epsilon_D; \epsilon_E] = 0$. To focus on the investment choice of workers in $t = 1$, q_j is assumed to be observed by workers and employers alike.

As is evident from the production functions, an important aspect of the model is that the first period effort of a worker contributes to the output of both periods. Hence, the first period is an "investment" period as well as a "test" period, which leads to an interpretation of e_1 as the amount of ability acquired through human capital investment. Consequently, it is natural to define "total" ability to be the sum of innate (a) and acquired (e_1) ability.

A key assumption made about the production technology is that $\alpha_E > \alpha_D (> 0)$ and $\gamma_D > \gamma_E (> 0)$. It implies that workers with lower ability have a comparative advantage in job E: Nonetheless, abler workers have absolute advantages in both jobs because output increases with ability, regardless of time period and job type.

The model so far is described in general terms to preserve a broad applicability to ordinary labor markets. Yet the model is particularly relevant to the issues of discrimination and affirmative action in college admission, and can be better understood in such a setting. In the latter case, $t = 1$ is a pre-college period when human capital investment is made, and q_{1i} is the test score of a college applicant i. Reinterpreting $j = D; E$ as "prestigious" and "less prestigious" colleges, respectively, q_j is then the output of a student that is associated with the attendance of a j-type college. The assumption regarding the second period production implies that a student who possesses a lower total ability (due to a lower innate ability and/or a lower investment in the pre-college period) is more suitable to a less selective college.

Assume that the distribution of innate ability is identical across groups and is normal: $a_i \sim N(\mu_0; \sigma^g)$, $g = B; W$. Workers are uncertain about their own innate ability, but are assumed to correctly know the first two moments of the distribution. Employers have a prior belief that the innate ability of a randomly drawn worker from group g is $\bar{\mu}_0^g$, with precision $1/\sigma^g$. This initial belief on the mean ability of g workers ($\bar{\mu}_0^g$) may or may not be equal to the true parameter μ_0, but the information about σ is precise. Employers do not observe the first period investment chosen by g workers, but conjecture that it is e_1^g. Then, from the perspective of firms, the conditional distribution of ability of a g worker after q_1 is observed is (DeGroot (1970)) normal with mean

$$\bar{\mu}_i^g \equiv E[a|q_{1i}; e_1^g; \bar{\mu}_0^g] = \frac{\sigma^g_{\epsilon_1}\bar{\mu}_0^g + \sigma^g(q_{1i} - e_1^g)}{\sigma^g_{\epsilon_1} + \sigma^g} \tag{2.4}$$

and variance $\frac{\sigma^g\sigma^g_{\epsilon_1}}{\sigma^g+\sigma^g_{\epsilon_1}}$. Note that the posterior belief $\bar{\mu}$ itself is distributed normally with mean $\bar{\mu}_0$ and variance $\frac{\sigma^2}{\sigma+\sigma_{\epsilon_1}}$ within each worker group. Denote this distribution by $H(\bar{\mu})$.[4]

To close the model, the wage schedules have to be specified. Consider the case where wage contracts are renegotiation proof in each period. Assume also that the wage contracts are linear: $w_1 = c_1 + b_1q_1$ in $t = 1$ and $w_j = c_j + b_jq_j$; $j = D; E$; in $t = 2$. In a college admission competition case, $t = 1$ is purely an investment period and $w_1 = 0$.

Optimal Job Assignment Rule

Start from $t = 2$. After e_1^* is chosen and the conditional expectation about his own innate ability is calculated using any information available (I), the expected utility of a worker is

$$c_j + b_j[\alpha_j + \gamma_j(E[a|I] + e_1^* + e_2)] - C(e_j); \tag{2.5}$$

[4] Since $a_i|q_{1i} \sim N(\bar{\mu}_i^g; \frac{\sigma^g\sigma^g_{\epsilon_1}}{\sigma^g+\sigma^g_{\epsilon_1}})$ and $\bar{\mu}_i^g \sim N(\bar{\mu}_0^g; \frac{(\sigma^g)^2}{\sigma^g+\sigma^g_{\epsilon_1}})$, the variance of ability after q_i's are observed is σ^g which is the same as the prior. This shows that the model is logically consistent.

should he be assigned to job j at $t = 2$. The optimal second period effort at j hence satisfies $b_j \gamma_j = C'(e_j^*)$. As can be expected from risk neutrality assumption, however, the optimal wage contract does not include "insurance" payments, i.e., $c_1 = c_j = 0$, so $b_1 = b_j = 1$ for $j = D, E$.[5] [6] Hence $e_D^* = C'^{-1}(\gamma_D) > e_E^* = C'^{-1}(\gamma_E)$. Note that the second period efforts depend only on the slope parameters of production technology.

Firms, conjecturing that g workers have chosen $\bar{e}_1^g$ so that the total ability is distributed around the mean $\bar{m}_0 + \bar{e}_1$ within that group, and knowing that any worker in job j will choose $e_j^*(\gamma_j)$ regardless of the group he belongs to, decide the job assignment rule: assign a worker to D if $\bar{m} > \hat{m}$, and to E, otherwise. The Pareto optimal contract specifies that $\hat{m}$ maximizes the net expected output of each group of workers which is given by

$$\int_{-\infty}^{\hat{m}} [\alpha_E + \gamma_E(\bar{m} + \bar{e}_1 + e_E^*) - C(e_E^*)]dH(\bar{m}) + \int_{\hat{m}}^{\infty} [\alpha_D + \gamma_D(\bar{m} + \bar{e}_1 + e_D^*) - C(e_D^*)]dH(\bar{m}).$$

Using an approximation $\gamma_j e_j^* - C(e_j^*) \doteq \gamma_j e_j^* - C'(e_j^*)e_j^* = 0$, the first order condition is, $\hat{m} + \bar{e}_1 \doteq \frac{\alpha_E - \alpha_D}{\gamma_D - \gamma_E} \equiv \theta$. This, combined with (2.4), implies that the optimal standard in terms of output is

$$\hat{q} = (1 + \rho)\theta - \rho(\bar{e}_1 + \bar{m}_0), \tag{2.6}$$

where $\rho \equiv \sigma_{\epsilon_1}/\sigma$.

---

[5] The proof is similar as in Gibbons and Murphy (1992). Free entry of firms implies that a firm's profit is zero. Furthermore, the profit from each job should also be zero; If a firm earns zero total profit while it gets positive profit from job D (E) and negative profit from E (D), then other firms can hire only D (E) workers and offer a slightly higher wage and make profits. Hence $c_E + b_E E[q_E|\bar{m} < \hat{m}] = E[q_E|\bar{m} < \hat{m}]$, implying that $c_E = (1 - b_E)E[q_E|\bar{m} < \hat{m}] = (1 - b_E)E[q_E|q_1 < \hat{q}]$, where $\hat{m}$ and $\hat{q}$ are the cut-off levels of ability and output chosen by firms as in (6). Likewise, $c_D = (1 - b_D)E[q_D|q > \hat{q}]$. From the perspective of firms, the expected utility of a worker who is assigned to D is

$$E[c_D + b_D(\alpha_D + \gamma_D(a + \bar{e}_1 + e_D^*(b_D\gamma_D)) + \epsilon_D|q_1 > \hat{q}] - C(e_D^*(b_D\gamma_D))$$

$$= E[\alpha_D + \gamma_D(a + \bar{e}_1 + e_D^*(b_D\gamma_D)) + \epsilon_D|q_1 > \hat{q}] - C(e_D^*(b_D\gamma_D)),$$

for an arbitrary b_D. An optimal contract chooses b_D that maximizes the above equation. Hence b_D^* satisfies $(\gamma_D - C'(e_D^*))\partial e_D^*/\partial b_D = 0$, which implies that $b_D = 1$ and $c_D = 0$. Likewise, $b_E = 1$ and $c_E = 0$. Since the expected utility from the perspective of a worker is the same as in the above equation except for a replacement of $\bar{e}_1$ by e_1^*, the contract is optimal for workers, too.

It can also be shown in a similar manner that the first period optimal contract specifies that $c_1 = 0$ and $b_1 = 1$.

[6] When workers do not observe q_1, firms may have an incentive to pay workers below their actual outputs. In that case, firms will make profits and workers will know that firms have cheated. I assume that firms behave honestly because they are liable for a breach of contract. Of course, unobservability of q_1 does not cause any problem in the case of college admission competition because $w_1 = 0$.

Equation (2.6) shows that the optimal output standard increases in θ and decreases in $\bar{e}_1$ and $\bar{m}_0$. A larger θ implies that the critical level of ability for which the expected output at two jobs are the same is larger. Naturally, the output standard will be higher when θ is larger. If, on the other hand, firms conjecture that workers in group W have acquired more human capital than group B workers ($\bar{e}_1^W > \bar{e}_1^B$), the optimal cut-off ability will be lower for W. This result is due to the fact that e_1 enters into the production in both periods, and, consequently, an increase in e_1 results in a greater productivity increase in D than in E.[7] Similarly, if employers believe that the mean innate ability of W workers is higher than that of B workers, they will "rationally" apply a lower output standard for the Ws. These standard lowering effects of $\bar{e}_1$ and $\bar{m}_0$ will be magnified when ρ is larger.[8]

Optimal Investment Decision

Taking the second period optimal contract and the output standard ($\hat{q}$) as given, a worker in $t = 1$ chooses an investment which maximizes the expected life-time utility

$$E[a + e_1 + \epsilon_1 - C(e_1)] + \beta \mathrm{Prob}(q_1 < \hat{q})\{E[q_E | q_1 < \hat{q}] - C(e_E^*)\}$$

$$+\beta \mathrm{Prob}(q_1 > \hat{q})\{E[q_D | q_1 > \hat{q}] - C(e_D^*)\}. \qquad (2.7)$$

Let $F(a + \epsilon_1; m_0)$ be the distribution function of $a + \epsilon_1 \sim N(m_0, \sigma + \sigma_{\epsilon_1})$ so $\mathrm{Prob}(q_1 < \hat{q}) = \mathrm{Prob}(a + \epsilon_1 < \hat{q} - e_1) = F(\hat{q} - e_1; m_0)$. Workers correctly know F.

The first order condition for the problem is (Appendix A)

$$1 + \beta[F(\hat{q} - e_1^*)\gamma_E + (1 - F(\hat{q} - e_1^*))\gamma_D] = C'(e_1^*). \qquad (2.8)$$

The interior solution is guaranteed by the assumption that $C' > 0$ for $e_1 \geq 0$. The equation states that the first period incentive is the sum of the direct incentive from the first period wage contract (which is zero when $t = 1$ is a pre-college investment period) and the implicit incentive from "career concerns" (Gibbons and Murphy (1992)). The latter

[7] In contrast, when e_1 has no effect on the second period production, the optimal output standard is given by $\bar{e}_1 + (1 + \rho)\theta - \rho\bar{m}_0$; so the standard is higher exactly by $\bar{e}_1$.

[8] Note also that an increase of ρ will result in a higher (lower) standard if $\bar{m}_0 + \bar{e}_1$ is smaller (larger) than θ. Hence, when employers believe that Bs' mean total ability is smaller than θ, they will rationally set a higher standard for B workers if Bs' "test scores" are subject to more noise (higher σ_{ϵ_1}) and/or B workers are more homogeneous (smaller σ).

is the discounted value of the weighted average of increased wages in the second period, where weights are the probabilities of being assigned to jobs D and E.

Equation (2.8) may have multiple solutions. To see this, modify the equation to

$$-F(\hat{q} - e_1) = k_1 C'(e_1) - k_2 \qquad (2.8)'.$$

Here $k_1 = 1/(\beta \Delta \circ)$ and $k_2 = (1 + \beta \gamma_D)/(\beta \Delta \circ)$; where $\Delta \circ \equiv (\gamma_D - \gamma_E)$. Clearly, $k_2 > 1$. Referring to Figure 1, the slope of $-F$ is greatest when $F = 1/2$, with its magnitude $f(m_0) = \frac{1}{\sqrt{2\pi(\sigma + \sigma_{\varepsilon_1})}}.$ Hence, (2.8) has three solutions if $f(m_0) > k_1 C'(e_1)$; i.e., if the marginal cost of investment and/or $\sigma + \sigma_{\varepsilon_1}$ are not large, ceteris paribus. Denote those solutions by $e_1^l; e_1^0$, and e_1^h. By the second order condition, e_1^0 is eliminated. Among two other solutions, e_1^h is the optimal investment level if the area under curve $-F$ between e_1^l and e_1^h is larger than the area under curve $k_1 C'(e_1) - k_2$ over the same region. Clearly there exists some $\hat{q}^0$ for which two areas are exactly the same. Assume that a worker in that case chooses e_1^h because of the "prestige" that is associated with D job. Then e_1^* is a decreasing function of $\hat{q}$ which is discontinuous at $\hat{q}^0$.

Comparative Statics

Define elasticities

$$\mu = \frac{e_1}{C'} \frac{dC'}{de_1}; \quad \eta = \frac{e_1}{\Delta_q} \frac{d\Delta_q}{de_1};$$

where Δ_q is the left hand side of (2.8). By the second order condition, $\mu > \eta$. Evaluating the elasticities at e_1^*, it is easy to establish (Appendix B) that

$$-\frac{\partial e_1^*}{\partial \hat{q}} = \frac{\partial e_1^*}{\partial m_0} = \frac{\eta}{\mu - \eta} > 0; \qquad (2.9)$$

confirming an earlier result that a worker facing a lower standard invests more. Also, when workers differ in their beliefs about their own innate abilities, those who (correctly or incorrectly) believe they have a higher innate ability invest more. For a given $\hat{q}$, such workers believe to have a higher chance of being assigned to D, and perceive that the expected marginal benefit of additional investment is greater. $e_1^*(\hat{q}; m_0)$ may be discontinuous at some values of its arguments.

Worker investment decreases with the marginal cost of investment. Suppose that $C(e_1) = \omega e_1^2/2$. Then

$$\frac{\partial e_1^*}{\partial \omega} = \frac{e_1^*}{\beta f(\hat{\theta} - e_1^*)\Delta\gamma - \omega} < 0. \tag{2.10}$$

Hence, if, for example, $\omega^B > \omega^W$, then $e_1^{*B} < e_1^{*W}$.

When the productivity differential between two jobs, $\Delta\gamma$, increases while $(\gamma_D + \gamma_E)/2$ remains the same,[9] its effect on investment depends on the magnitude of mean total ability relative to $\hat{\theta}$:

$$\frac{\partial e_1^*}{\partial \Delta\gamma} = -\frac{\beta(1/2 - F(\hat{\theta} - e_1^*))}{\beta f(\hat{\theta} - e_1^*)\Delta\gamma - C''(e_1^*)} \geq (<) 0 \text{ if } m_0 + e_1^* \geq (<) \hat{\theta}. \tag{2.11}$$

Intuitively, when a worker's total ability is greater than the output standard, his chance of being assigned to D exceeds 1/2. Consequently, an increase in the relative productivity in D gives more incentive to invest.

$e_1^*(\hat{\theta})$, with other arguments suppressed, summarizes workers' best responses to firms' job assignment policy. It is discontinuous at some value of its argument when ω and/or $\sigma + \sigma_{\varepsilon_1}$ are sufficiently small.

3. The Market Equilibrium

The market equilibrium is implicitly determined by the interaction between (2.6) and (2.8). The assignment rule (2.6) is the firms' best policy for a given set of prior information about the ability of workers. Note that what determines firms' policy in (2.6) is their belief about the mean of the sum of innate and acquired ability of workers, $\hat{x}(\equiv \hat{m}_0 + \hat{e}_1)$. Hence rewrite (2.6) as

$$\hat{\theta} = (1 + \rho)\theta - \rho\hat{x}. \tag{3.1}$$

The actual magnitude of mean total ability, $x(\equiv m_0 + e_1^*)$; is implicitly a function of $\hat{x}$ via its effect on $\hat{\theta}$

$$x(\hat{x}) = m_0 + e_1^*(\hat{\theta}(\hat{x})); \tag{3.2}$$

which describes workers' best response to the firms' policy.

[9] This may well characterize the 1980s when the relative return to higher skill increased.

It is then natural to define the market equilibrium as a pair of $(x^*; \tilde{x}^*)$ which satisfies

$$x^*(\tilde{x}^*) = \tilde{x}^*. \tag{3.3}$$

Under the assumption that a and ϵ_1 are normally distributed, the distribution of q_1 is completely described by its first two moments. Since σ and σ_{ϵ_1} are assumed to be known, firms' beliefs about the distribution of workers' test scores are confirmed in market equilibrium. That is, the market outcome is belief-consistent. Note that, when firms have a biased belief about the mean innate ability of workers belonging to a group, their conjecture on the investment level of those workers is not equal to the actual investment in market equilibrium. For example, $\tilde{e}^B > e_1^{*B}$ when $\tilde{m}_0^B < m_0$. The equilibrium condition requires that firms' beliefs should be consistent not with the mean innate or acquired ability of workers but with the sum of them.

There exists at least one solution to (3.3).[10] Of particular interest to the current work, however, is if there are multiple equilibria. Clearly, an increase in $\tilde{x}$ can result in a faster increase of $x(\tilde{x})$ if $e_1^*(\hat{q}(\tilde{x}))$ is very responsive to the change in the output standard over some range of $\tilde{x}$. In that case, $x(\tilde{x})$ can cut the $x = \tilde{x}$ line from below, resulting in three equilibria. Assume that employers' beliefs about the mean ability of g workers in generation $\tau + 1$ equals the actual ability of the generation τ; $\tilde{x}^g_{\tau+1} = x^g_\tau$. Then, the equilibrium in the middle is unstable. Recall that $e_1^*(\hat{q}(\tilde{x}))$ is discontinuous at some $\tilde{x}'$ when $\beta \cdot f(m_0) > C'$, which increases the chance of getting multiple equilibria. Since $e_1^*(\hat{q}(\tilde{x}))$ in (3.2) is the solution to (2.8) which involves a cumulative normal distribution, an analytical solution is difficult to get. Still, it can be shown (Appendix C) that multiple equilibria exist when ω and/or $\sigma + \sigma_{\epsilon_1}$ are small.

Refer to the $x(\tilde{x})^1$ curve in Figure 2. It is drawn for case 1 with the properties that it is discontinuous at some point, the derivative of $x(\tilde{x}) - \tilde{x}$ has two solutions, and the marginal costs of investment are identical across groups. It has two stable equilibria at

[10] The proof is as follows. Note first that $e_1^*(\hat{q}(\tilde{x})) > 0$ for $\tilde{x} > -1$ so $x|_{\tilde{x}=0} = m_0 + e_1^*(\hat{q}(0)) > 0$. Furthermore, the slope of $x(\tilde{x})$, $\frac{\partial e_1^*}{\partial \hat{q}}\frac{\partial \hat{q}}{\partial \tilde{x}} = \rho\frac{\eta}{1-\eta} > 0$ approaches to zero as $\tilde{x}$ increases indefinitely. That is, $x(\tilde{x})$ in a $(\tilde{x}; x(\tilde{x}))$ plane has a positive intercept and has a slope which eventually becomes zero. Therefore, $x(\tilde{x})$ intersects with $x = \tilde{x}$ (from above) at least once.

$\tilde{x}^B$ and $\tilde{x}^W$. Both are belief-consistent or self-fulfilling equilibria. Those equilibria are discriminatory because the lower ex post ability distribution of B workers in equilibrium is entirely due to the biased belief of employers about Bs' innate ability. Note also that a large difference in the ex post abilities across groups can be obtained not only by a small difference in the employers' initial belief about the mean innate ability of workers but also by a mild taste for discrimination.

When discriminatory equilibria exist, employers have no incentive to deviate once one of the equilibria is attained. This must become clearer with the following discussion. The sequence of actions implicit in the process is that firms choose output standards first and workers choose investment later. Employers' beliefs initiate all subsequent actions and outcomes. Thus, a crucial aspect of the model is that the workers' response summarized by (3.3) is not taken into account by employers. Assume that workers' investments do not change when an employer changes his output standard. Suppose that there is an employer who knows the whole market process and has no negative stereotypes on the ability of B workers. If he decides to lower the standard for group B, he will be assigning more underqualified workers to job D than his competitors, and will have a lower output. Hence, he has no incentive to change his action. The same is true when every employer knows the whole market process: the coordination problem lies in the way.[11]

Based on Phelps' (1972) hypothesis that traditional indicators of ability for workers in certain "disadvantaged" groups are less informative, Lundberg and Startz (1983) show that those workers have less incentive to make productivity-enhancing investments in human capital. When discriminatory equilibria exist, a similar result is obtained. Consider a case in which groups W and B are identical except that B workers' test scores are noisier than their counterpart's: $\rho^B > \rho^W$. Appendix D shows that the output standard for group B will be even higher, which in turn discourages their investment even more. Therefore, when discriminatory equilibria exist, noisier test scores magnify the difference of average wages as well as investments between W and B workers. This demonstrates that the

[11] One might think that employers have an incentive to experiment with various $\hat{y}$ in an attempt to find out true e_1^* and m_0. It is not the case, however, because of the reason given above.

current model provides a unified framework in which both of Arrow-type and Phelps-type statistical discrimination can be analyzed.

Now consider case 2 in which groups differ in investment costs. Suppose that $\omega^B > \omega^W$.[12] Then, as discussed earlier, $e_1^{*B} < e_1^{*W}$ for all possible $\bar{x}$, so $x(0)^B < x(0)^W$. Furthermore, the slope of $x(\bar{x})$, $\rho\beta\Delta \cdot f(\hat{\theta} - e_1)/\{\omega - \beta\Delta \cdot f(\hat{\theta} - e_1)\}$, is decreasing in ω. It follows that the slope and intercept of Bs' $x(\bar{x})$ will be smaller than those of Ws. Also, should the curve have a jump, it occurs at a smaller $\bar{x}$ than for group W. Suppose that $x(\bar{x})^B$ still has two equilibria. Clearly, $x^{*B} < x^{*W}$. Now a part of the difference in mean abilities between groups is due to the difference in the marginal costs of investment. Statistical discrimination only magnifies ability differences.

Finally, imagine case 3 in which ω^B and ω^W are such that the $x(\bar{x})^B$ curve, $x(\bar{x})^2$ in Figure 2, has only one intersection with $x(\bar{x}) = \bar{x}$ line, while $x(\bar{x})^W$, as $x(\bar{x})^1$ in the figure, has multiple intersections.[13] In this case, even though there is no room for statistical discrimination, B workers' mean ability is still substantially lower than that of W workers. This shows that a small difference in the marginal costs of investment across groups can result in a large difference in the ability distributions even in the absence of any statistical discrimination.[14]

4. Affirmative Action and Mismatch

Statistical Discrimination and Wage Differentials

Consider two equilibria $x^*(\bar{x}^g) = \bar{x}^g$ in Figure 2 and their corresponding output stan-

[12] The marginal cost of financing investment may be higher for Bs because they are from families with lower incomes (which may be partly due to discrimination in the past). Also, Bs may be in an environment which is not very favorable for human capital investments. In that case, the investment efficiency, which is defined to be the amount of acquired ability for a given expenditure of investment cost, may be lower for them. Clearly, the group differential in the marginal costs of investment captures that kind of investment efficiency differential.

[13] This case is also obtained when Bs believe (correctly or incorrectly) that their innate ability is lower than that of Ws.

[14] There may be only one equilibrium for both groups if ω is large and/or $\sigma + \sigma_{\epsilon_1}$ is very large. The current paper is interested in the instances where discriminatory equilibria exist, and does not pursue the single equilibrium case further.

dards $\hat{q}^g(\hat{x}^g)$, $g = B, W$.[15] To simplify the discussion, take the case where σ and σ_{ε_1} are identical across groups. The average wages (outputs) of g workers in D and E are

$$\hat{q}^g_D = \alpha_D + \gamma_D(e_1^{*g} + e_D^{*} + m_0 + \kappa\lambda^+(d^g)); \ \hat{q}^g_E = \alpha_E + \gamma_E(e_1^{*g} + e_E^{*} + m_0 - \kappa\lambda(d^g));$$

respectively. Here $\lambda^+(d^g) = \frac{\phi(d^g)}{1-\Phi(d^g)}$; $\lambda(d^g) = \frac{\phi(d^g)}{\Phi(d^g)}$; $\kappa \equiv \frac{\sigma}{\sqrt{\sigma+\sigma_{\varepsilon_1}}}$; $d^g = \frac{\hat{q}^g - x^{*g}}{\sqrt{\sigma+\sigma_{\varepsilon_1}}}$; and ϕ and Φ are the pdf and cdf of the standard normal distribution, respectively.

The wage differential (or, test score differential in the college admission example) between B and W in $t = 1$ is $\Delta q_1 = e_1^{*B} - e_1^{*W}$. The wage differentials in D and E are

$$\Delta\hat{q}_D = \gamma_D[(e_1^{*B} - e_1^{*W}) + \kappa(\lambda^+(d^B) - \lambda^+(d^W))](< 0)$$

$$\Delta\hat{q}_E = \gamma_E[(e_1^{*B} - e_1^{*W}) - \kappa(\lambda(d^B) - \lambda(d^W))](< 0); \qquad (4.1)$$

respectively. The first term in the square bracket of each equation, which is negative, is the wage difference that reflects an investment differential between two groups. The second term reflects sample selections. Because B workers face a higher output standard, the average innate abilities of B workers in both D and E are higher than those of W workers, which mitigates the wage differentials in both jobs. However, as Appendix E shows, the net effect of statistical discrimination on the average wage of B workers is negative in both jobs.

Normalize the population size of each generation to 1 and let the share of B workers in the population in each generation be $s(< 1/2)$. When investment costs are identical across groups (case 1), the aggregate net output of group B in $t = 2$ is lower than that attainable in the absence of discrimination by

$$\Delta q_2 = s\Delta\gamma\{(x^W - x^B)\gamma_D/\Delta\gamma + \hat{F}^W\cdot(\theta - x^W + \kappa\lambda(d^W)) - \hat{F}^B\cdot(\theta - x^B + \kappa\lambda(d^B))\}; \qquad (4.2)$$

where $\hat{F}^g \equiv F(\hat{q}^g - e_1^{*g})$. That magnitude is the second period aggregate welfare loss which is entirely due to statistical discrimination. But the welfare loss is smaller when the marginal cost of investment is higher for B workers. For example, statistical discrimination is directly responsible for none of the wage differentials in case 3.

[15] The following analysis works with any stable equilibrium.

Affirmative Action

Observing that B workers, compared to W workers, are assigned to a high paying job (a prestigious college) less often and also receive lower average wages in both periods and both jobs, a regulatory agency might ascribe such a disparity to the "unfair" practice of employers and attempt to correct it by enforcing affirmative action. However, as is clear from the above discussion, the economic achievement of group B will be lower than that of W not only when discriminatory equilibria exist but also when Bs' marginal investment cost is higher. Affirmative action might not enhance efficiency in the latter case.

Theoretically, a regulatory agency can adopt two different types of policy in enforcing affirmative action. One policy would ensure that workers with the same test scores are treated equally in job assignment. The other policy would enforce an equal representation of workers across groups in job D. However, as pointed out by Coate and Loury (1993 b), enforcing an equal standard requires that the regulator can observe all information used by employers in making assignment decisions. Also, as Lundberg (1991) argues, employers can evade regulations when making assignment decisions by using other variables which are correlated with group characteristics. Furthermore, the outcome of such a policy is qualitatively identical to that of the second type of policy in the current model.

By focusing on the economic effects of the second type of affirmative action policy, and supposing that the policy is enforced (unexpectedly) after workers have already made their investments, the problem of an employer is to find out standard q^g which maximizes the net expected output of workers. The employers think that $Prob(q_1 > q^g) = 1 - F(q^g - e_1^g; x^g; \sigma + \sigma_{\epsilon_1})$ proportion of workers will be assigned to D if q^g is adopted. The affirmative action constraint that the proportion of B workers assigned to D is no smaller than that of W workers implies that $F(q^B - e_1^B) = F(q^W - e_1^W)$, because employers have no reason to assign a higher proportion of B workers to D. Hence, an employer's problem is

$$\max_{q^B;q^W} s[F(q^B - e_1^B)\{E[q_E|q_1 < q^B] - C(e_E^*)\} + (1 - F(q^B - e_1^B))\{E[q_D|q_1 > q^B] - C(e_D^*)\}]$$

$$+(1 - s)[F(q^W - e_1^W)\{E[q_E|q_1 < q^W] - C(e_E^*)\} + (1 - F(q^W - e_1^W))\{E[q_D|q_1 > q^W] - C(e_D^*)\}] \quad (4.3)$$

subject to the equality constraint.

The constraint implies that $q^B - q^W = -(\bar{x}^W - \bar{x}^B) = -(e_1^{*W} - e_1^{*B}) < 0$, which in turn implies that the standard differential should be exactly the same as the opposite of the investment differential between groups. Ignoring $\gamma_j e_j^* - C(e_j^*)$, the solution is

$$q^g = (1+\rho)\theta - \rho x^g; \tag{4.4}$$

where $x^B \equiv \{1 + (1-s)(1+1/\rho)(\bar{x}^W/\bar{x}^B - 1)\}\bar{x}^B$ and $x^W \equiv \{1 - s(1+1/\rho)(1 - \bar{x}^B/\bar{x}^W)\}\bar{x}^W$.[16]

Equation (4.4) states that, in order to comply with the affirmative action, employers have to adopt standards that would have been chosen if the mean ability of g workers was actually x^g. Note that x^B (x^W) is decreasing (increasing) in s. Clearly, $x^B > \bar{x}^B$ and $x^W < \bar{x}^W$. Also, since $q^B < q^W$, employers have to behave as if they believe that the mean ability of group Bs is higher than that of W. Furthermore, it can be verified that $x^B > \bar{x}^W$ (implying $q^B < \hat{q}^W$) if $s < 1/(1+\rho)$. The latter condition is satisfied if s is sufficiently small and/or ρ is not very large (e.g., $\rho < 1$). In such a case, employers choose the standard for group B as if its mean ability is higher than the actual mean ability of group W.[17]

Suppose that workers can observe q_1. Then, similarly as in (2.4), a worker's posterior belief about his own innate ability after observing q_{1i} is $E[a|q_{1i}; e_1^{*g}; m_0] = \frac{\sigma_{\varepsilon_1} m_0 + \sigma(q_{1i} - e_1^{*g})}{\sigma_{\varepsilon_1} + \sigma}$. Hence, ignoring $\gamma_j e_j^* - C(e_j^*)$, the critical level of output ($\hat{q}^*$) on which a worker's acceptance decision will be based satisfies $\alpha_E + \gamma_E(E[a|\hat{q}^{*g}; e_1^{*g}; m_0] + e_1^*) = \alpha_D + \gamma_D(E[a|\hat{q}^{*g}; e_1^{*g}; m_0] + e_1^*)$, yielding $\hat{q}^{*g} = \hat{q}^g$. Therefore, once investments are already made, workers will use exactly the same output standards that would have been used by employers in the absence of affirmative action. This kind of self-selection of workers does not cause any problem to employers in recruiting W workers to job D; because more Ws are willing to take job D

[16] It can be easily checked that the second order condition is satisfied.

[17] The intuition is as follows. Facing an equal representation constraint, employers have to balance the "output loss" from assigning overqualified Ws to E job and that from assigning underqualified Bs to D job. Since the total output loss increases in the number of workers in each group, the optimal response of employers when s is small is to raise the standard for Ws (Bs) a little (much). If s is sufficiently small, then only Bs' standard has to be lowered while Ws' standard is kept almost intact.

than are offered. But among $1 - F(q^B - e_1^B)$ portion of B workers with job D offers, only $1 - F(\hat{q}^B - e_1^{*B})$ portion will accept them. This shows that when q_1 is observed by workers employers will not be able to comply with affirmative action at least for the generation to which the action is applied for the first time.

Alternatively, consider a case in which workers cannot observe q_1 while employers can. For example, in a college admission competition case where q_1 is associated not only with the absolute level of "test scores" of an individual but also with his rank among the pool of applicants to a college, an individual may have far less accurate information about his rank than the college. Suppose that workers still have a good idea about the standard. A worker then knows that his test score is no smaller than q^g when he is offered a D job. It follows that his posterior belief of ability is $E[a|q_1 > q^g] = m_0 + \kappa\lambda^+(d^g)$; where $d^g = \frac{\hat{q}^g - x^{*g}}{\sqrt{\sigma^2 + \sigma^2_{\epsilon_1}}}$. Hence, the condition that all Bs who are offered D jobs will accept them is that $x^{*B} - \theta + \kappa\lambda^+(d^B) \geq 0$.[18] Since $\lambda^+(d^B)$ increases with s, employers can comply with affirmative action only when s is not very small and/or $\theta - x^{*B}$ is not large.

Mismatch

Continue to take a case where workers cannot verify q_1 and the equal representation constraint is also compatible with workers' self-selection. As noted earlier, when an investment e_1^{*g} is already made, the optimal cut-off test score is $\hat{q}^g$. Those standards ensure that workers utilize comparative advantages and consequently guarantee the maximum (net) total output. Under affirmative action, however, employers have to set standards $q^g (\neq \hat{q}^g)$. Since the B workers with test score $q_1 \geq \hat{q}^B$ ($q_1 \leq q^B$) and W workers with $q_1 \geq q^W$ ($q_1 \leq \hat{q}^W$) will be assigned to D (E) in any case, the action changes the job assignment of the Bs with $q_1 \in [q^B; \hat{q}^B)$ and the Ws with $q_1 \in [\hat{q}^W; q^W)$. Those workers are mismatched in that they are assigned to jobs in which they do not have comparative advantages. Such a mismatch is additional to the one that is inevitable due to the uncertainty regarding the innate abilities of workers.

For the B workers with $q_1 \in [q^B; \hat{q}^B)$ who are under-qualified for Ds but are assigned

[18] A worker will accept a D job offer only if $\alpha_D + \gamma_D(E[a|q_1 > q^g] + e_1^*) \geq \alpha_E + \gamma_E(E[a|q_1 > q^g] + e_1^*)$, which yields the condition.

to them, the expected net output in D is

$$E[\alpha_D + \gamma_D(a + e_1^{*B}) | q^B < q_1 < \hat{q}^B] = \alpha_D + \gamma_D\left(x^{*B} - \kappa \frac{\phi(d^B) - \phi(d^B)}{\Phi(d^B) - \Phi(d^B)}\right). \qquad (4.5)$$

Those workers could have produced net output amounting to

$$\alpha_E + \gamma_E\left(x^{*B} - \kappa \frac{\phi(d^B) - \phi(d^B)}{\Phi(d^B) - \Phi(d^B)}\right), \qquad (4.6)$$

were they assigned to E. Since the number of B workers who are mismatched is $s(\hat{F}^B - F^B)$, where $F^B \equiv F(q^B - e_1^{*B})$, etc., the (net) output loss due to the misallocation of B workers is

$$\Delta q^B = s \cdot \gamma \{ \hat{F}^B \cdot (\theta - x^{*B} + \kappa \lambda(d^B)) - F^B \cdot (\theta - x^{*B} + \kappa \lambda(d^B)) \} > 0. \qquad (4.7)$$

Likewise, the total (net) output loss due to the assignment of over-qualified Ws to E is

$$\Delta q^W = (1 - s) \cdot \gamma \{ \hat{F}^W \cdot (\theta - x^{*W} + \kappa \lambda(d^W)) - F^W \cdot (\theta - x^{*W} + \kappa \lambda(d^W)) \} > 0. \qquad (4.8)$$

The total net output loss in the second period is $\Delta q^B + \Delta q^W$, which is the sum of the output losses due to the "favoritism" toward group B and the "reverse discrimination" toward group W.

Comparing two regimes with and without affirmative action, which one yields a smaller output loss? To answer this question, it is necessary to specify transitional dynamics that are dependent upon the expectations and adjustments in behavior of employers and workers. Continue to assume that employers adjust their beliefs about the ability of g workers according to $x^g_{\tau+1} = x^g_\tau (= x^{*g}_\tau)$. Correspondingly, workers in generation $\tau + 1$ know that the standard for them will be $(1 + \rho)\theta - \rho x^{*g}_\tau$.

Consider first a case in which investment costs are identical across groups. The welfare loss due to mismatch in the presence of affirmative action can be shown to always be smaller than that which arises from discrimination in the absence of the policy: $\Delta q_2 > \Delta q^W + \Delta q^B$. Furthermore, once affirmative action is enforced, the investment differential eventually disappears, and, consequently, economic achievements become identical across groups.

The reason is as follows. Note that q^B which corresponds to x^B in Figure 2 induces the second generation B workers to have the mean ability $x^*(x^B)$ $(> x^*(x^W))$:[19] Hence, unless prohibited by the action, workers in group B will be promoted to D more often than W workers. Employers, "surprised" to observe that the mean ability of group B is $x^*(x^B)$, will then choose standard $\hat{q}^B(x^*(x^B))$ $(> q^B)$ for the third generation of B workers, reducing their investment. This process will continue until x^B converges to the point $x^*(\bar{x}^W) = \bar{x}^W$. This shows that an enforcement of affirmative action for one generation will eventually eliminate statistical discrimination in the market forever if groups do not differ in investment costs. Also, the output gain attained by eliminating discrimination outweighs the output loss due to mismatch that appears during the transitional period. Therefore, (a temporary introduction of) affirmative action enhances efficiency and also guarantees identical productivity across groups in the long run.

When the costs of investment differ across groups, however, affirmative action does not necessarily enhance efficiency. Consider case 3. When employers choose $q(x^B)$ to comply with the action, the second generation of workers will have the mean ability $x^{*'}(x^B)$ in Figure 2. It is higher than that of previous generation B workers, $x^*(\bar{x}'^B)$: Observing that the mean ability of group B is now $x^{*'}(x^B)$, employers will be able to comply with the action even when they raise the standard for the next generation, which reduces the investment incentive of B workers in that generation. Assuming $s < 1/(1+\rho)$; this process will continue until x^B_τ converges to some point which is located at the right of $\bar{x}^W$: In this "long run" equilibrium, the mean ability of group B is still lower than that of W. Hence, B workers will have to be treated preferentially while "reverse discrimination" will continue to be the case for group W. The mismatch effect will never be wiped out, and affirmative action policy will neither eliminate negative stereotypes of employers about B workers[20]

[19] Note that this is true regardless of observability of q_1. The only difference in transitional dynamics when q_1 is observed by workers is that employers cannot comply with the action for the first generation to which affirmative action is applied. On the other hand, if a regulatory agency announces in advance that the affirmative action will be introduced in the next generation, the response of workers in the first generation will be the same as that of the second generation described in the text.

[20] In Coate and Loury (1993 b), affirmative action may not eliminate negative stereotypes of employers when the fraction of qualified workers increases in standard for some range. In the current model, worker investment always decreases in standard, and, consequently, the proportion of workers whose test scores

nor guarantee equal outputs across groups. Also, equal representation of workers across groups in job D will be attained only when the policy is a permanent institution in the market: the equilibrium returns to $x^*(\bar{x}'^B) = \bar{x}'^B$ if the policy is abolished.

An Application

The model gives insight into why the drop-out rates and the returns to college education differ among blacks between those who attended historically black colleges (HBCs) and those who attended integrated colleges. As stated earlier, the first period in the college admission competition case is the pre-college period when students produce "test scores" upon which college admissions are based. The second period productivity parameters are then the returns associated with the performance of students in the prestigious and less prestigious colleges. The model suggests that minority students who are admitted to prestigious colleges with the help of affirmative action (e.g., minority quota, etc.) may perform more poorly there than at less prestigious colleges, and consequently are more likely to drop out of those colleges.[21] In addition, since performance in the college, measured by GPA, is a more important determinant of earnings than the prestigiousness of the college (James, Alsalam, Conaty, and To (1989)), those students may have lower earnings.[22]

There is some evidence supporting these predictions. The drop-out rate among blacks who attend HBCs is almost the same as that of whites in all 4 year colleges, but blacks who attend non-HBCs drop out twice as often as their counterpart's in HBCs (Hoffman, Snyder, and Sonnenberg (1992)). Steele (1992) reports that in a prestigious college from 18 to 33 percent of black students flunked out while only from 2 to 11 percent of whites flunked out. The average letter grade of black students was also much lower than that of

exceed the standard always decreases in Δ.

[21] Affirmative action in the college admission case may have purposes other than to correct for the unjust outcome arising from statistical discrimination. Whatever those purposes are, the mismatch described in the text will continue to be the consequence of the action, as long as the Bs in equilibrium have acquired less human capital in pre-college years.

[22] It can be verified that the average output in D is higher than that in E by $\Delta\delta[x^{*B} - \theta + \eta\lambda^+(d^B)] + \lambda^+(d^B)\gamma_E = F^B + \gamma_D e_D^* - \gamma_E e_E^*$. It is positive because the square bracket is positive by the self-selection compatibility condition. Hence, strictly speaking, the model does not necessarily implies that the Bs in D on average produce less output than the Bs in E. When the model is extended to a case in which workers (students) derive a substantial amount of psychic income from getting D jobs, however, it is easy to show that there is a case in which almost everybody with D offers accept them and the output in D is lower than that in E.

whites. On the other hand, Loury and Garman (1993) report that even though blacks gain more in earnings by attending selective colleges when performance is held constant, their gains are o®set by lower performance. Furthermore, Constantine (1994) ¯nds that the returns from college education are higher for blacks who attended HBCs than for blacks who attended non-HBCs. In sum, this evidence may indicate that the mismatch e®ect induced by a±rmative action is substantial.

Discussion

The economic consequences of a±rmative action are analyzed in this paper under the assumption that preference-based discrimination as in Becker (1957) is absent in the market. As noted earlier, however, discriminatory equilibria can be obtained even when employers initially have a very mild taste for discrimination. Preference-based discrimination exercised once can result in a permanent economic disparity across demographic groups because of the self-perpetuating nature of those discriminatory equilibria. Consequently, the analysis can be readily extended to the case where employers have a taste for discrimination.

Suppose that there is a continuum of jobs which can be ranked by the complexity of task and productivity as there are numerous colleges of di®ering \qualities". In such a case, a±rmative action can generate a chain reaction of mismatch. Workers who ¯t the second most di±cult job best are assigned to the most di±cult job, which forces employers to assign workers who best ¯t the third or fourth most di±cult jobs to the second most di±cult job, and so on. The mismatch will be widespread, and consequently the economic ine±ciency will be substantial, at least during the transitional period.

The outcome is similar when there are many stages in promotion competition. Suppose that there are multiple periods, T; in one's working life, a worker faces promotion competition as he advances to each consecutive period. Since workers who are treated preferentially in promotion decision will perform more poorly than they could have performed in the absence of such treatments, their chance of being quali¯ed for higher level jobs in later periods will be smaller, too. The equal representation constraint will be binding more stringently for both groups in later periods. Therefore, the output losses from mismatch will be larger for workers in higher level jobs in later periods. Such losses will occur permanently unless investment costs are identical across groups.

Suppose that workers differ in their innate abilities. Suppose further that a is still normally distributed across workers but each worker has an unbiased belief about his own innate ability: As is clear from the discussion in Section 2, a worker with a higher a will invest more. Now it is difficult to characterize equilibria because $a + e_1^*(a)$ may no longer be normally distributed due to the discontinuity of $e_1^*(a)$: Nonetheless, the analysis will be unchanged as long as employers adjust their beliefs about the total ability of a median worker using the median of the first period output.

As noted earlier, affirmative action eventually leads to economic efficiency if the investment costs are identical across groups. However, when there is a substantial difference in such costs (case 3), no output difference reflects an inefficiency due to discrimination. Affirmative action in that case induces B workers to increase investment, but at the expense of matching efficiency, and equal economic achievement across groups is not obtained. Actually, in stark contrast with the case of females, the academic achievement of blacks measured by SAT score, enrollments in law and medical schools, etc., has not improved dramatically, even after the introduction of affirmative action. This may indicate that the blacks indeed face a substantially higher cost of investment in human capital. If it is really the case, an alternative policy that can guarantee identical productivity across groups is the one which gives a subsidy to B workers for their investments. A policy which gives lump-sum money to the workers in group B who are successful in getting job D could have similar results. Such policies will be efficient if the subsidy is financed through lump-sum taxes on all workers. But to maintain efficiency through this policy the tax-subsidy should remain in permanence.[23]

[23] Since the subsidy given to a generation induces the workers in that generation to invest more and receive higher wages, the tax-subsidy policy may not need to be a permanent fixture if the marginal cost of investment becomes lower for subsequent generations through some kind of inter-generational transmission mechanism.

5. Conclusion

Even though a±rmative action is a very important institution governing labor markets, not much e®ort has been made to assess its economic consequences. This paper identi¯es the mechanism through which a±rmative action results in a mismatch between workers of di®ering abilities and tasks of di®ering complexities, by incorporating statistical discrimination, workers' human capital investment incentives, comparative advantages of workers in jobs of di®ering complexities, and a±rmative action in a uni¯ed framework. It then compares the welfare loss due to discrimination in the absence of a±rmative action policy and the welfare loss due to mismatch in the presence of the policy. The e±ciency of the policy is shown to depend much upon the investment cost di®erentials across demographic groups.

This framework provides a new insight into some stylized facts that have not been explained before. For example, the theory, when reinterpreted appropriately, may explain why drop-out rates have been lower, but earnings have not, for blacks who attended historically black colleges when compared to those who enrolled in non-black colleges. The model also suggests that previous research which attempted to assess the economic e®ects of a±rmative action may have exaggerated the gains of minority workers by ignoring the fact that the action induces a misallocation of workers.

Appendix

A. Derivation of equation (2.8)

The second term in (2.7) is

$$\mathrm{Prob}(q_1 < \hat{q})E[q_E | q_1 < \hat{q}] = F(\hat{q} - e_1)\{\alpha_E + \gamma_E(E[a|q_1 < \hat{q}] + e_1 + e_E^*)\}$$

$$= F(\hat{q} - e_1)\{\alpha_E + \gamma_E(e_1 + e_E^*)\} + \gamma_E \int_{-\infty}^{\hat{q}-e_1} \frac{\sigma^2_{\varepsilon_1} m_0 + \sigma^2(a + \varepsilon_1)}{\sigma^2_{\varepsilon_1} + \sigma^2} dF(a + \varepsilon_1).$$

Manipulating the expected output in D in a similar manner, the expected utility of a worker at $t = 1$ can be written as

$$m_0 - e_1 - C(e_E^*) + \beta\Big[F(\hat{q} - e_1)\{\alpha_E + \gamma_E(e_1 + e_E^*) - C(e_E^*)\} + \gamma_E \int_{-\infty}^{\hat{q}-e_1} \frac{\sigma^2_{\varepsilon_1} m_0 + \sigma^2(a + \varepsilon_1)}{\sigma^2_{\varepsilon_1} + \sigma^2} dF(a+\varepsilon_1)\Big]$$

$$+\beta\Big[(1 - F(\hat{q} - e_1))\{\alpha_D + \gamma_D(e_1 + e_D^*) - C(e_D^*)\} + \gamma_D \int_{\hat{q}-e_1}^{\infty} \frac{\sigma^2_{\varepsilon_1} m_0 + \sigma^2(a + \varepsilon_1)}{\sigma^2_{\varepsilon_1} + \sigma^2} dF(a + \varepsilon_1)\Big].$$

The first order condition is,

$$1 - C'(e_1^*) + \beta\{F(\hat{q} - e_1^*)\gamma_E + (1 - F(\hat{q} - e_1^*))\gamma_D\}$$

$$-\beta f(\hat{q} - e_1)[\{\alpha_D + \gamma_D(m(\hat{q}; e_1^*) + e_1^* + e_D^*) - C(e_D^*)\} - \{\alpha_E + \gamma_E(m(\hat{q}; e_1^*) + e_1^* + e_E^*) - C(e_E^*)\}];$$

where $m(\hat{q}; e_1^*) = \frac{\sigma^2_{\varepsilon_1} m_0 + \sigma^2(\hat{q} - e_1^*)}{\sigma^2_{\varepsilon_1} + \sigma^2}$. The expression in the square bracket is the difference in the expected utilities between jobs D and E when $q_1 = \hat{q}$, and is zero, because the expression in the square bracket $= \alpha_D - \alpha_E + (\gamma_D - \gamma_E)[m(\hat{q}; e_1^*) + e_1^*] + \gamma_D e_D^* + \gamma_E e_E^* - C(e_D^*) + C(e_E^*) \doteq \alpha_D - \alpha_E + (\gamma_D - \gamma_E)[m(\hat{q}; e_1^*) + e_1^*] = 0$.

B. Derivation of (2.9)

The derivation of $\partial e_1^*/\partial \hat{q}$ is straightforward. For $\partial e_1^*/\partial m_0$; (2.8) yields

$$\frac{\partial e_1^*}{\partial m_0} = \frac{\beta \cdot \Delta \cdot [\partial F(\hat{q} - e_1^*)/\partial m_0]}{\beta f(\hat{q} - e_1^*) \cdot \Delta - C''(e_1^*)}.$$

Note that

$$\frac{\partial F(\hat{q} - e_1^*)}{\partial m_0} = \frac{\partial}{\partial m_0} \int_{-\infty}^{\hat{q}-e_1^*} \frac{1}{\sqrt{2\pi(\sigma^2 + \sigma^2_{\varepsilon_1})}} \exp\Big\{-\frac{(a + \varepsilon_1 - m_0)^2}{2(\sigma^2 + \sigma^2_{\varepsilon_1})}\Big\} d(a + \varepsilon_1)$$

$$= \frac{1}{\sigma^2 + \sigma^2_{\varepsilon_1}} \int_{-\infty}^{\hat{q}-e_1^*} \frac{1}{\sqrt{2\pi(\sigma^2 + \sigma^2_{\varepsilon_1})}} (a + \varepsilon_1 - m_0) \exp\Big\{-\frac{(a + \varepsilon_1 - m_0)^2}{2(\sigma^2 + \sigma^2_{\varepsilon_1})}\Big\} d(a + \varepsilon_1)$$

$$= \frac{1}{\sigma + \sigma_{\varepsilon_1}}\left\{\int_{-\infty}^{\hat{q}-e_1^*}(a+\varepsilon_1)g(a+\varepsilon_1)d(a+\varepsilon_1) - m_0F(\hat{q}-e_1^*)\right\}$$

$$= \frac{1}{\sigma + \sigma_{\varepsilon_1}}F(\hat{q}-e_1^*)\left\{E[a+\varepsilon_1|a+\varepsilon_1<\hat{q}-e_1^*] - m_0\right\}.$$

Since $E[a+\varepsilon_1|a+\varepsilon_1<\hat{q}-e_1^*] = m_0 - \sqrt{\sigma+\sigma_{\varepsilon_1}}\,\lambda(d)$, where $\lambda(d) = \phi(d)/\Phi(d)$, ϕ and Φ are the pdf and cdf of the standard normal distribution, respectively, and $d = (\hat{q}-e_1^*-m_0)/\sqrt{\sigma+\sigma_{\varepsilon_1}}$,

$$\frac{\partial F(\hat{q}-e_1^*)}{\partial m_0} = -\frac{F(\hat{q}-e_1^*)}{\sqrt{\sigma+\sigma_{\varepsilon_1}}}\lambda(d) < 0.$$

Hence,

$$\frac{\partial e_1^*}{\partial m_0} = -\frac{\beta F(\hat{q}-e_1^*)\cdot\gamma\lambda(d)}{\sqrt{\sigma+\sigma_{\varepsilon_1}}\{\beta f(\hat{q}-e_1^*)\cdot\gamma - C''(e_1^*)\}} > 0.$$

The inequality follows from the S.O.C. for (2.7). Dividing both numerator and denominator by $\beta f(\hat{q}-e_1^*)\cdot\gamma$ and using the definitions of elasticities,

$$\frac{\partial e_1^*}{\partial m_0} = \frac{1}{\sqrt{\sigma+\sigma_{\varepsilon_1}}}\frac{F(\hat{q}-e_1^*)}{f(\hat{q}-e_1^*)}\frac{\eta}{\mu-\eta}\lambda(d) = \frac{1}{\sqrt{\sigma+\sigma_{\varepsilon_1}}}\frac{\eta}{\mu-\eta}\frac{F(\hat{q}-e_1^*)}{\Phi(d)}\frac{\phi(d)}{f(\hat{q}-e_1^*)}$$

$$= \frac{\eta}{\mu-\eta};$$

where the last equality is due to the fact that $\phi(d)/\sqrt{\sigma+\sigma_{\varepsilon_1}} = f(\hat{q}-e_1^*)$ and $F(\hat{q}-e_1^*) = \Phi(d)$.

C. Existence of Multiple Equilibria

The possibility of existence of multiple equilibria is explored here in a case where $C(e_1) = \frac{\omega e_1^2}{2}$. Since $C'(e_1) = \omega e_1$ and $C''(e_1) = \omega$, $(2.8)'$ becomes

$$-F(\hat{q}-e_1) = k_1'e_1 - k_2;$$

where $k_1' = \omega/(\beta\cdot\gamma)$ and k_2 is the same as in $(2.8)'$. Hence, the RHS of the equation is a straight line. Referring to Figure 1, the condition for an existence of three solutions to equation $(2.8)'$ is $f(m_0) > k_1'$, or, equivalently, $\frac{\beta\cdot\gamma}{\omega\sqrt{\sigma+\sigma_{\varepsilon_1}}} > \sqrt{2\pi}$.

Let $\hat{q}^0$ be the one for which $-F$ cuts $k_1'e_1 - k_2$ line from below at $F = 1/2$, and its corresponding three solutions be $e_1^l; e_1^0$, and e_1^h. Because f is symmetric, a worker is indifferent between e_1^l and e_1^h, in which case he is assumed to choose e_1^h. Clearly, $e_1^*(\hat{q})$

is discontinuous at $\hat{q}^0$ because, $e_1^* < e_1^l$ if $\hat{q} > \hat{q}^0$, and $e_1^* \geq e_1^h$; otherwise. At e_1^0, $-F = -1/2 = k_1^{\prime} e_1 - k_2$, implying that $e_1^0 = \frac{1}{2k_1^{\prime}}(2k_2 - 1) = (1 + \beta\Delta\gamma/2)/\omega.$ Also, since m_0 satisfies $m_0 = \hat{q}^0 - e_1^0$, it follows that $\hat{q}^0 = (1+\rho)\theta - \rho\tilde{x}^0 = m_0 + e_1^0$. Hence, $x(\tilde{x})$ curve is discontinuous at $\tilde{x}^0 = (1 + 1/\rho)\theta - (1 + \beta\Delta\gamma/2)/(\rho\omega) - m_0/\rho.$ One implication of this result is that $\tilde{x}^0$ is smaller for a group with a larger ω. Note also in passing that the slope of $x(\tilde{x})$ is greatest at $\hat{q}^0(\tilde{x}^0) - e_1(\hat{q}^0(\tilde{x}^0)) = m_0$. Hence, if $x(\tilde{x})$ curve is discontinuous, then it is at the inflection point of $x(\tilde{x})$.

Referring to $x(\tilde{x})^1$ in Figure 2, multiple market equilibria exist when $x(\tilde{x}^0) = m_0 + e_1^h > \tilde{x}^0$ and $x(\tilde{x}^{0-}) = m_0 + e_1^l < \tilde{x}^0$, where $x(\tilde{x}^{0-}) = \lim_{\varepsilon \to 0}(\tilde{x}^0 + \varepsilon)$. This condition is clearly satisfied if $m_0 + e_1^0 = \tilde{x}^0$, or, equivalently, $m_0 = \theta - (1 + \beta\Delta\gamma/2)/\omega.$ More generally, when $x(\tilde{x})$ has a jump, there will be multiple equilibria if $e_1^l < \tilde{x}^0 - m_0 < e_1^h$, or, equivalently, $e_1^l < (1 + 1/\rho)\theta - (1 + \beta\Delta\gamma/2)/(\rho\omega) - (1 + 1/\rho)m_0 < e_1^h$. Since $x(\tilde{x}^0) - x(\tilde{x}^{0-}) = e_1^h - e_1^l = \beta\Delta\gamma[F(\hat{q}^0 - e_1^l) - F(\hat{q}^0 - e_1^h)]/\omega$ is decreasing in ω and $\sigma + \sigma_{\varepsilon_1}$, the chance of getting two equilibria increases as the marginal cost of investment and/or $\sigma + \sigma_{\varepsilon_1}$ gets smaller, for given $m_0; \theta; \rho$, and $\beta\Delta\gamma$.

D. Negative Impacts of Noisier Test Scores on Bs' Investments

From (3.1), $\frac{\partial \hat{q}}{\partial \rho} = \rho(\theta - \tilde{x})$; which implies that a larger ρ results in a higher $\hat{q}$ for Bs and lower $\hat{q}$ for Ws if $\tilde{x}^B < \theta < \tilde{x}^W$. Consider the equilibrium $x(\tilde{x}^B) = x^{*B} = \tilde{x}^B$ in Figure 2. The corresponding output standard satisfies, $\hat{q}^B = (1+\rho)\theta - \rho x^{*B} = (1+\rho)\theta - \rho(m_0 + e_1^{*B}) > (1+\rho)\theta - \rho(m_0 + e_1^0) = (1+\rho)\theta - \rho\hat{q}^0.$ The inequality follows from $e_1^0 > e_1^l > e_1^B$; and the last equality is due to the fact that $\hat{q}_0 - e_1^0(\hat{q}^0) = m_0$ (see Appendix C). Hence, $\hat{q}^B + \rho\hat{q}^0 > (1+\rho)\theta.$ Since $\hat{q}^B > \hat{q}^0$, it follows that $\hat{q}^B > \theta$; which in turn implies that $\hat{q}^B - \theta = \rho(\theta - \tilde{x}^B) > 0.$ This shows that, when discriminatory equilibria exist, $\tilde{x}^B$ is always smaller than $\theta.$ Therefore, an increase of ρ due to noisier test scores (larger σ_{ε_1}) results in a higher standard, and reduces the investments of Bs via (3.2).

E. The Signs in (4.1)

Equation (2.6) implies that $\tilde{x}^B - \tilde{x}^W = (\hat{q}^W - \hat{q}^B)/\rho.$ But from Figure 2, $\tilde{x}^B - \tilde{x}^W =$

$x^*(\hat{x}^B) - x^*(\hat{x}^W) = e_1^{*B} - e_1^{*W}$. Hence $e_1^{*B} - e_1^{*W} = (\hat{q}^W - \hat{q}^B)/\rho$. Note further that

$$\lambda^+(d^B) - \lambda^+(d^W) \doteq \frac{d\lambda^+(\bar{d})}{dd} \cdot (d^B - d^W);$$

where $\bar{d} = (d^B + d^W)/2$. Since $1 \geq \frac{d\lambda^+(\bar{d})}{dd} = \lambda^+(\bar{d}) \cdot (\lambda^+(\bar{d}) - \bar{d})) \geq 0$, and $d^B - d^W = \frac{1}{\sqrt{\sigma+\sigma_1^2}}(\hat{q}^B - e_1^{*B} - \hat{q}^W + e_1^{*W}) = -\frac{1+\rho}{\sqrt{\sigma+\sigma_1^2}}(e_1^{*B} - e_1^{*W})$;

$$\Delta\bar{q}_D \doteq \gamma_D(e_1^{*B} - e_1^{*W})(1 - \lambda^+(\bar{d}) \cdot (\lambda^+(\bar{d}) - \bar{d})) < 0;$$

and

$$\Delta\bar{q}_E \doteq \gamma_E(e_1^{*B} - e_1^{*W})(1 - \lambda(\bar{d}) \cdot (\lambda(\bar{d}) + \bar{d})) < 0;$$

which establishes the result.

References

[1] Aigner, Dennis J. and Cain, Glenn G. \Statistical Theories of Discrimination in Labor Markets" Industrial and Labor Relations Review. 30 (January 1977): 175-87.

[2] Arrow, Kenneth J. \The Theory of Discrimination" in Orley Ashenfelter and Albert Rees, eds, Discrimination in Labor Markets. Princeton, NJ: Princeton University Press, 1973, pp. 3-33.

[3] Becker, Gary S. The Economics of Discrimination, Chicago, IL: University of Chicago Press, 1957.

[4] Borjas, George J., and Goldberg, Mathews S. \Biased Screening and Discrimination in the Labor Market." Amer Econ Rev. 68 (December 1978): 918-22.

[5] Coate, Stephen, and Loury, Glenn C. \Antidiscrimination Enforcement and the Problem of Patronization." Amer Econ Rev Papers and Proceedings 83 (May 1993 a): 92-8.

[6] Coate, Stephen, and Loury, Glenn C. \Will A±rmative-Action Policies Eliminate Negative Stereotypes?" Amer Econ Rev 83 (December 1993 b): 1220-40.

[7] James, Estelle, Nabeel Alsalam, Joseph C. Conaty, and Duc-Le To. \College Quality and Future Earnings: Where Should You Send Your Child to College?" Amer Econ Rev Papers and Proceedings 79 (May 1989): 247-52.

[8] Constantine, Jill M. \Measuring the E®ect of Attending Historically Black Colleges and Universities on Future Wages of Black Students." mimeo, Williams College, 1994.

[9] Daniel, Kermit. \Self-Selection, Endogenous Wage O®ers and Statistical Discrimination." mimeo, University of Pennsylvania, October 1993.

[10] Gibbons, Robert, and Murphy, Kevin J. \Optimal Incentive Contracts in the Presence of Career Concerns." J Pol Econ 100 (June 1992): 468-505.

[11] Ho®man, Charlene M., Snyder, Thomas, and Sonnenberg, Bill. Historically Black Colleges and Universities, 1976-1990. Washington, DC: National Center for Education Statistics, 1992.

[12] Loury, Linda D., and Garman David. \A±rmative Action in Higher Education." Amer Econ Rev Proceedings 83 (May 1993): 99-103.

[13] Kahn, Lawrence M. \Customer Discrimination and A±rmative Action." Econ Inquiry 26 (July 1991): 555-71.

[14] Lang, Kevin. \A Language Theory of Discrimination." Quarterly J Econ 101 (May 1986): 363-82.

[15] Leonard, Jonathan. \The impact of A±rmative Action on Employment." J Labor Econ 2 (October 1984): 439-63.

[16] Lundberg, Shelly J. \The Enforcement of Equal Opportunity Laws Under Imperfect Information: A±rmative Action and Alternatives." Quarterly J Econ 106 (February 1991): 309-26.

[17] Lundberg, Shelly J., and Startz, Richard. \Private Discrimination and Social Intervention in Competitive Labor Markets." Amer Econ Rev 83 (June 1983): 340-7.

[18] Milgrom, Paul, and Oster, Sharon. \Job Discrimination, Market Forces, and the Invisibility Hypothesis." Quarterly J Econ 102 (August 1987): 453-76.

[19] Phelps, Edmund S. \The Statistical Theory of Racism and Sexism." Amer Econ Rev 62 (September 1972): 659-61.

[20] Rosen, Sherwin. \Substitution and Division of Labor." Economica 45 (August 1978): 235-50.

[21] Schotter, Andrew, and Weigelt, Keith. \Asymmetric Tournaments, Equal Opportunity Laws, and A±rmative Action: Some Experimental Results." Quarterly J Econ 107 (May 1992): 511-39.

[22] Schwab, Stewart. \Is Statistical Discrimination E±cient?" Amer Econ Rev 76 (March 1986): 228-34.

[23] Smith, James P., and Welch, Finis. \A±rmative Action and Labor Markets." J Labor Econ 2 (April 1984): 269-301.

[24] Steele, Claude M. \Race and the Schooling of Black Americans." The Atlantic Monthly (April 1992): 68-78, reprinted in Feiner Susan F., eds., Race and Gender in the American Economy, Englewood Cli®s, NJ: Prentice Hall, 1994, pp. 172-82.

[25] Welch, Finis. \Employment Quotas for Minorities." J Pol Econ 84 (August 1976): S105-39.

[26] Welch, Finis. \A±rmative Action and Discrimination." in Stephen Shulman and William Darity, Jr., eds., The Question of Discrimination, Middletown, CT: Wesleyan University Press, 1989, pp. 153-89.