Statistical Discrimination and Intergenerational Mobility in a Matching Model

Peter Arcidiacono*

December 22, 1999

Abstract

I develop a matching model with statistical discrimination and intergenerational mobility. Workers make both observable and unobservable investments before entering the market which affect their future productivity. Firms can search for workers based upon observable characteristics, be it observable investments, age, or some exogenous characteristic such as race. Multiple equilibria within each market can exist, one in which workers make unobservable investments and many firms search and one in which workers do not make unobservable investments and few firms search. Hence, two groups of workers that differ on an observable, exogenous characteristic (say, race) can be in two different equilibria. If parents' investments decisions affect the investment decisions of their children, policies which remove the statistical discrimination by pooling the black and white labor markets will still lead to unequal results in the short run even if the groups are *ex ante* identical. Empirical predictions on the relationship between experience, education, and wages result.

Keywords: Discrimination, search, intergenerational mobility
JEL D31, J40, J70.

---

*Department of Economics, Duke University. Psarcidi@econ.duke.edu. The author thanks Kate Antonovics, John Kennan, Jonathan Parker and Marc Rysman for helpful comments as well as the Alfred P. Sloan Foundation for financial assistance.

1 Introduction

Wage inequality across the races, as well as the programs designed to decrease it, have been a topic of much discussion in both the popular press and among economists. The debate among economists has been focused on how wage inequality can exist even if the groups are *ex ante* identical. The literature has focused on two types of discrimination: taste discrimination, where owners (or possibly customers and other workers) receive disutility from hiring workers of a particular type,[1] and statistical discrimination, where profit maximizing firms make rational decisions under uncertainty which lead to discriminatory outcomes.[2]

One strain of the statistical discrimination literature was begun by Coate and Loury (1993). Here, statistical discrimination is where owners (correctly) infer that workers of a particular type have not, on average, made an unobservable investment which influences their productivity. Workers make an investment with firms observing only a noisy signal of whether an investment was made. Firms use an observable feature (say race) in helping to infer whether or not the workers have made investments. Coate and Loury show that an equilibrium exists where whites tend to invest and are assigned jobs where the investment is useful and blacks do not invest and are never assigned the good jobs. What distinguishes this from a model where firms do not want to hire blacks because they do not like them is that here firms do not care about skin color in and of itself, they care only about maximizing profits. In their version of affirmative action, a patronizing equilibrium exists where blacks do not invest and are still assigned the good jobs.[3]

Since Coate and Loury, there have been two papers which are relevant to the work here. Antonovics (1999) extends Coate and Loury's model to allow for intergenerational transfers. She shows that discriminatory outcomes can persist even after the implementation of an equal opportunity law. Mailath, Samuelson, and Shaked (1998) get discriminatory outcomes from a different source: the search intensities of firms. In their model there are two sectors; one where a wage is fixed and the other which depends critically on one's ability to match with a firm. Their model has firms searching harder for workers of a particular type which

[1] See Becker (1957)

[2] See Aigner and Cain (1977), Arrow (1973), Lundberg and Startz (1983), and Phelps (1972) for earlier works on statistical discrimination.

[3] Chung (1998) shows how to get around the patronizing equilibrium using implementation theory. Moro and Norman (1999) apply a variant of Coate and Loury's model and show how the model works in a general equilibrium context.

leads to the discriminatory outcome.

This paper uses a simple matching model which generates many of the attractive features of these last two papers- similar to Antonovics it produces discriminatory outcomes even when workers who are identical except for the observable feature are treated the same and similar to Mailath, Samuelson, and Shaked in that the tool of the statistical discrimination is how firms search for workers. However, this paper is not limited to producing their results in a simple manner. The simplicity of the model allows for the addition of observable investment decisions and gains to experience. With these additions, many empirical predictions result regarding how statistical discrimination affects the interaction between education, experience, and wages.

The model is similar to Pissarides (1992) in that it is an overlapping generations matching model where workers live for tow periods each. When young, workers have an opportunity to make an investment which enhances their productivity. There are two types of workers, say b types and w types. Expected productivity across types is the same before any investments have been made. Firms and workers search for each other, with firms being able to dictate which type of workers they want to search for. Matched young workers see their match dissolve with some probability s and have to search anew if their match dissolves. All wages are determined through generalized Nash bargaining and workers in preserved matches renegotiate their wages when old.

The model is able to generate multiple equilibria because the unobservable investment decision and the number of searching firms both depend positively on one another. In particular, workers are more likely to make the investment the higher the probability is that they will match with a firm. The probability of matching with a firm is increasing in the number of searching firms. However, the number of searching firms is also increasing in the number of workers who make the investment. Hence, multiple equilibria result where both no investment and investment can be supported as equilibrium outcomes. The model then naturally leads to one in which there is statistical discrimination and firms search for workers on the basis of some observable feature, say skin color. One group would then be in the 'no investment' equilibrium while the other group would be in the 'investment' equilibrium. I show under fairly general conditions that workers in the no investment equilibirium have flatter wage profiles with less variance within a particular generation than their counterparts in the investment equilibrium.

Wage inequality across the groups can be eliminated within one generation through an equal opportunity law which requires firms to search equally hard for each type of worker. However, if the cost of investing depends upon decision of one's parent, wage inequality will not disappear immediately.[4] In particular, if the cost of investment is affected by the parent's decision to invest in such a way that there is positive state dependence in the decision to invest, wage inequality across types will take time to disappear. Further, in the short run, the present value of wages for the workers who were in the investment equilibrium fall as the probability of finding a match is lower than in the equilibrium with investment.[5] Over the course of time, the wage profiles for the discriminated group will become steeper and the variance of wages within a particular cohort will increase.

Since the decision to invest depends upon the decision of one's parent, it is clear how an economy that had slavery could lead into an economy with statistical discrimination. With slavery, blacks had no incentive to invest.[6] Hence, even if there was only one equilibrium when the economy has the stationary distribution of skills associated with the investment equlibrium, this is not where blacks began. Multiple equilibria can then arise through the initial conditions of the groups. Blacks could then be permanently trapped in the no investment equilibrium.

With the discriminatory outcome not immediately removed, other programs would need to be implemented to immediately eliminate wage inequality across types. I define two classes of programs: equal opportunity, where firms are mandated to treat workers who are identical except for type the same, and affirmative action, where groups receive subsidies or special employment opportunities based upon skin color. One such affirmative action program which would accelerate the removal the discriminatory outcome would be to subsidize a portion of the investment cost for the discriminated group. In fact, I show that w types (the nondiscriminated group) may find contributing to such a program attractive initially, with less of a subsidy being provided for each subsequent generation. By subsidizing the cost of investment for some of the b types, more firms will enter which helps w types as well. There may reach a point, however, when w types no longer wish to contribute and that this point

[4] See Antonovics (1999) for another model of how quickly statistical discrimination can be remedied.

[5] See Moro and Norman for another model where removing statistical discrimination has a negative effect on the group that was not being discriminated against.

[6] Blacks also may not have an incentive to invest under taste discrimination and this too could create the state dependence.

may occur before the wage distributions converge.

The model is then expanded to allow for observable investment decisions, education being one example. Firms are allowed to search based upon whether a worker has made observable investments. Hence, separate labor markets exist for each observable investment category.[7] I show that statistical discrimination still holds with the degree of the wage penalty for b types being higher in the observable investment market.

One of the features of this literature is that the models are too cumbersome to generate meaningful empirical predictions about the relationship between statistical discrimination and such things as experience and education. The model here has implications for the wage profiles of the two groups of workers where statistical discrimination exists and how these wage profiles change after an equal opportunity law is implemented. The predictions for the wage profiles could be tested empirically, treating statistical discrimination as the case before the 1964 Civil Rights Act and treating the economy as one where an equal opportunity law has been implemented after.

The rest of the paper proceeds as follows. Section 2 presents the basic model. Section 3 shows how the model can generate multiple equilibria if workers make an investment in their skills before entering the workforce. This leads to statistical discrimination which can be corrected by an equal opportunity law. Section 4 extends the statistical discrimination to allow the investment decisions of the parents to affect the investment decisions of the children. Section 5 investigates the effects of an affirmative action program where b type workers have their investment subsidized. Section 6 adds observable investment decisions to the model. Empirical predictions of the model are provided in section 7. Section 8 concludes.

2 The Model

In this section I present an overlapping generations matching model where workers live for two periods. There is no population growth: I workers are born each period. The timing of decisions for a particular cohort follows:

1. All young workers search for firms. Matched young workers and firms negotiate a wage.

[7] This model could easily be extended to one where 'separate labor markets' is replaced by 'search intensities.'

2. When old, workers who were employed see their match dissolve with probability s. All preserved matches renegotiate their wage.

3. Old workers who are not in a preserved match search. Those who find matches then negotiate a wage.

2.1 The Matching Technology

All unemployed workers search. Workers search for firms and firms search for workers. Firms can search on observables which in this case is age.[8] The number of searching firms for young (old) workers, J_1 (J_2), is endogenous and determined by a zero profit condition on entry. If a worker matches when young, the match dissolves when old with probability s. Hence, the number of searching workers in the young market, I_{u1}, include all young workers. The number of searching workers in the old market, I_{u2}, include old workers who had their matches dissolve and old workers who were unable to find a match when young.

The total number of new filled jobs in market m, $m \in \{1, 2\}$ is a function of the number of searching firms and workers and is defined by:

$$x_m = \min\{x(J_m, I_{um}), J_m, I_{um}\} \tag{1}$$

where $x(J_m, I_{um})$ has positive partial derivatives, is concave in both J_m and I_{um}, and exhibits constant returns to scale.

Each worker (firm) has the same probability as any other worker (firm) of finding a job. Let the probability of a searching worker in the mth market finding a match be given by q_m. Hence, $q_m = \frac{x(J_m, I_{um})}{I_{um}}$. The corresponding probability for a searching firm in the mth market is then $\frac{x(J_m, I_{um})}{J_m}$.

2.2 Production

Production is a function of θ_i, an individual specific component which is exogenous[9] and distributed $G(\theta)$ where $\theta \in [0, \overline{\theta}]$. If a worker matches in the first period, he acquires human capital which is useful for the second period. In particular, experienced workers see their productivity enhanced by an additional $\gamma_e \theta_i$ where $\gamma_e \in \Re_+$. Let d_i indicate whether the

[8] Later we will add to the observables skin color as well as observable investments.

[9] In section 3, workers are allowed to make an investment decision which affects their draw on θ.

ith worker was employed in the previous period. Production for a worker firm pair, $\{i, j\}$, is then given by:

$$f_{ij} = (1 + d_i \gamma_e)\theta_i \tag{2}$$

2.3 Wages

Workers and firms negotiate a wage at the begining of each period. This negotiation in each period takes place even among preserved matches. Hence, in the young market workers negotiate over the current value of the match plus the change in the worker's outside option should the match dissolve. The bargaining power of the worker is set at β, $\beta \in (0, 1)$. The outside option for workers is zero for both tenured and untenured workers because at the time of the negotiations the search phase is complete.

Firms pay a fixed cost of entry, k_m, to search in the mth labor market. Firms must pay this cost each time they want to search for a new worker. Firms are restricted only by the matching function as to how many workers they can hire. Hence, the firm's outside option for not hiring a particular worker is zero. Let 1 and 2 indicate whether the worker is young or old. Wages for the ith worker then follow:

$$w_{i2} = \beta(1 + d_i \gamma_e)\theta_i \tag{3}$$

$$w_{i1} = \beta\theta_i - \frac{(1-\beta)\beta s q_2 \gamma_e \theta_i}{1+r} \tag{4}$$

where r is the interest rate, q_2 is the probability of a worker finding a match in the old market, and d_i is again whether the worker was employed in the previous period. The final term for young workers is negative as the worker not only accumulates skills that are useful at the particular firm, but useful elsewhere as well. Since the worker and the firm are Nash bargaining over the current period surplus of the contract as well as the change in the outside options of the worker, included in this is the gain that the worker receives from having skills that can be used elsewhere.

2.4 Equilibrium

I now show that, given values for all parts of the problem except for k_1 and k_2, for any positive values of k_1 and k_2 an equilibrium exists. An equilibrium is defined as a pair $\{J_1, J_2\}$ such that expected profits in market m equal zero if $J_m > 0$ and are negative otherwise. The expected zero profit conditions on entry into the two markets then define the equilibrium. The expected profit functions can be written as:

$$E(\pi_1) = \frac{IE(\theta)q_1(1-\beta)}{J_1}\left(1+\frac{1}{1+r}\left[(1-s)(1+\gamma_e)+s\beta\gamma_e q_2\right]\right)-k_1 \tag{5}$$

$$E(\pi_2) = \frac{IE(\theta)q_2(1-\beta)}{J_2}[sq_1(1+\gamma_e)+(1-q_1)]-k_2 \tag{6}$$

The expected profit function for entering the young market depends upon the probability of finding a match plus the current surplus of that match, plus the probability that the match stays times the surplus of a preserved match, plus the probability the match breaks up times the gain the worker receives from having matched in the previous period. For the old market, the expected profit function depends upon the probability of finding a match times the surplus associated with matching with a formerly employed or never employed worker times the probability of matching with a formerly employed or never employed worker, respectively.

Suppose an interior solution exists. That is, when $x_m = x(J_m, I_{um})$. In this case, we can solve for J_m as a function of q_m. Note that $q_m = x\left(\frac{J_m}{I_{um}}, 1\right)$. Define f as the inverse function of x. Then $I_{um}f(q_m) = J_m$. Both the first and second derivatives of f_m are positive as x_m is increasing and concave in J_m.[10] Expected profits in the two markets can then be written as:

$$E(\pi_1) = \frac{E(\theta)(1-\beta)q_1}{f(q_1)}\left(1+\frac{1}{1+r}\left[(1-s)(1+\gamma_e)+s\beta\gamma_e q_2\right]\right)-k_1 \tag{7}$$

$$E(\pi_2) = \frac{E(\theta)(1-\beta)q_2}{[q_1 s+(1-q_1)]f(q_2)}[sq_1(1+\gamma_e)+(1-q_1)]-k_2 \tag{8}$$

[10] Here I prove that $f(q_m)$ is increasing and convex in q_m. Let $z_m = \frac{J_m}{I_u m}$. Holding I_{um} constant, x_m is increasing in z_m. Hence, $\frac{\partial z_m}{\partial q_m} = \frac{1}{\frac{\partial q_m}{\partial z_m}} > 0$, which implies that $\frac{\partial q_m}{\partial z_m} > 0$. Differentiating again with respect to q_m, we have:

$$\frac{\partial^2 z_m}{\partial q_m^2} = -\frac{1}{\left(\frac{\partial q_m}{\partial z_m}\right)^2}\frac{\partial\left(\frac{\partial q_m}{\partial z_m}\right)}{\partial q_m}\frac{\partial^2 z_m}{\partial q_m^2} = -\frac{1}{\left(\frac{\partial q_m}{\partial z_m}\right)^2}\frac{\partial\left(\frac{\partial q_m}{\partial z_m}\right)}{\partial z_m}\frac{\partial z_m}{\partial q_m}\frac{\partial^2 z_m}{\partial q_m^2} = -\frac{1}{\left(\frac{\partial q_m}{\partial z_m}\right)^3}\frac{\partial^2 q_m}{\partial z_m^2}$$

Since we have already established that $\frac{\partial q_m}{\partial z_m} > 0$, the second derivative of the inverse function is the opposite sign of the second derivative of the function itself. This implies that the inverse function is convex in q_m. QED.

Let m' index the market that is not market m (m' is the old market if m is the young market, etc.). The derivatives of expected profits in market m with respect to q_m holding $q_{m'}$ constant are:

$$\begin{aligned}
\frac{\partial E(\pi_1)}{\partial q_1}|_{q_2} &= \frac{E(\theta)(1-\beta)q_1}{f(q_1)^2}\left[f(q_1) - q_1\frac{\partial f(q_1)}{\partial q_1}\right] \\
&\times\left(1+\frac{1}{1+r}[(1-s)(1+\gamma_e)+s\beta\gamma_e q_2]\right) \qquad (9) \\
\frac{\partial E(\pi_2)}{\partial q_2}|_{q_1} &= \frac{E(\theta)(1-\beta)q_2}{f(q_2)^2}\left[f(q_2) - q_2\frac{\partial f(q_2)}{\partial q_2}\right]\left(1+\frac{s\gamma_e q_1}{q_1 s+(1-q_1)}\right) \qquad (10)
\end{aligned}$$

Note that the terms in brackets are both negative,[11] implying that the sign of $\frac{\partial E(\pi_m)}{\partial q_m}|q_{m'}$ is negative. Hence, entry in market m, which increases q_m, leads to lower profits in market m all else equal.

However, the derivatives of expected profits in market m with respect to $q_{m'}$ holding q_m constant are both positive.

$$\begin{aligned}
\frac{\partial E(\pi_1)}{\partial q_2}|_{q_1} &= \frac{E(\theta)(1-\beta)q_1}{f_1(q_1)}\left(\frac{s\beta\gamma_e}{1+r}\right) \geq 0 \qquad (11) \\
\frac{\partial E(\pi_2)}{\partial q_1}|_{q_2} &= \frac{E(\theta)(1-\beta)q_2}{f_2(q_2)}\frac{s\gamma_e}{[q_1 s+(1-q_1)]^2} \geq 0 \qquad (12)
\end{aligned}$$

The higher the probability of matching in market m, which is increasing in the number of firms in market m, the higher the profits in market m'; there is a positive feedback between entry in market m and entry in market m'. If the market for young workers becomes tighter (high q_1), then the probability of a firm in the old market finding a trained worker increases. Since trained workers are more valuable to firms, q_2 also increases. Firms in the young labor market benefit from a tight labor market for old workers because this increases their bargaining position with young workers.

I now show that an equilibrium exists for all pairs of k_1 and k_2.

[11] Proof: Note that by the intermediate value theorem we have:

$$\frac{f_m(q_m) - f_m(0)}{q_m} = \frac{\partial f_m(\zeta)}{\partial \zeta}$$

for some $\zeta \in [0, q_m]$. Since $f_m(0) = 0$, we can rewrite the above equation as:

$$f_m(q_m) = q_m\frac{\partial f_m(\zeta)}{\partial \zeta} \leq q_m\frac{\partial f_m(q_m)}{\partial q_m}$$

where the last inequality comes from the convexity of f_m in q_m; the first derivatives of f_m are increasing in q_m. QED.

Theorem 2.1 *Given I, s, r, β, γ_e, x, and $G(\theta)$ there is pair $\{J_1, J_2\}$ that satisfies the two zero profit conditions for all positive values of k_1 and k_2.*

Proof: See the Appendix.

3 Pre-Market Investment and Statistical Discrimination

Up until now, the θ's, that is the productivities of the workers, have been taken as exogenous. I now let workers have some control over their θ's. The model then produces multiple equilibria. The multiple equilibria make statistical discrimination possible. The statistical discrimination can be eliminated through an equal opportunity law though the short term consequences are lower wages for the group that was not being discriminated against. Unless otherwise noted, all proofs are in the Appendix.

I now allow young workers to influence their productivities (their θ's) before entering the labor market. In particular, let the initial productivities of the workers, which are redefined as θ''s, be drawn from a distribution $G(\theta')$. If workers do not make an investment then their initial productivities equal their actual productivities. That is, $\theta_i' = \theta_i$. However, if workers do make an investment their productivity is $(1+\gamma_u)\theta_i'$, where $\gamma_u \in \Re_+$. Since $\theta_i' \geq 0$, with no costs all workers would make the investment. However, workers must pay a cost c to make the investment. I assume that workers do not know their productivities until after the investment decision is made. The following inequality then defines whether a worker will make the investment.

$$\beta\gamma_u\left(q_1 + \frac{1}{1+r}\Big[q_1(1-s)(1+\gamma_e) + sq_1q_2(1+\beta\gamma_e) + (1-q_1)q_2\Big]\right) \geq c \qquad (13)$$

The expression then holds with equality for some cutoff c^*.

Recall that q_m is increasing in $q_{m'}$. Because of this positive feedback, it is possible in equilibrium for profits to be increasing in q_m as higher values of q_m lead to higher values of $q_{m'}$. I rule out this case with the following assumption:

Assumption 3.1: *The matching function is such that, at the equilibrium values of q_m and $q_{m'}$, $-\frac{\partial E(\pi_m)}{\partial q_m}|q_{m'} > \frac{\partial E(\pi_m)}{\partial q_{m'}}|q_m$.*

Assumption 3.1 means that the positive feedback between the two markets does not outweigh the increased costs of having more firms in market m and hence a higher q_m.[12] Without

[12] A sufficient condition for this assumption to hold is that the matching function is $J_m^{\alpha} I_{um}^{1-\alpha}$ where $\alpha \leq \frac{1}{2}$.

Assumption 3.1, there is potentially an equilibrium where the probabilities of matching in each market are *decreasing* in $E(\theta)$. This would be the case when increased entry has little effect on the probability of matching, but the increase in profits in market m from entry in market m' is large. However, there would also be an equilibrium where $q_1 = q_2 = 1$.[13] Hence, Assumption 3.1 means that profits are competed away through increased entry rather than through exit.

I am now in a position to analyze how the investment decision and the probabilities of matching in each market depend upon one another. Note that the individual takes the probability of matching in each market as given. The following lemma shows how the reservation cost changes with changes in q_1 and q_2.

Lemma 3.1 *c^* is increasing in q_1 and q_2.*

The higher q_1 and q_2 are, the higher the probability the worker will reap a return on his investment and the more the worker is willing to pay for the investment. However, the probability of matching depends upon the proportion of people making the investment.

Lemma 3.2 *q_1 and q_2 are both increasing in $E(\theta)$.*

Since θ represents how productive the worker is, firms would obviously prefer to have workers with high values for θ. If $E(\theta)$ increases, more firms will need to enter for the zero profit condition to hold.

With θ having a positive effect on q_1 and q_2, and both q_1 and q_2 having a positive effect on the number of workers making the investment, the model may have multiple equilibria. In particular,

Theorem 3.1 *It is possible to set c such that at least two equilibria exist: one in which no workers invest and one in which all workers invest.*

This model naturally leads to one in which there is statistical discrimination. Consider the case where there are two types of workers, b and w. Both types of workers face the same c and the same distributions for their initial productivities. Now, let firms choose to search for either workers of b type or of w type.[14] This is the key assumption: statistical discrimination

[13] See Pissarides (1992).

[14] Statistical discrimination from firms being able to search on worker type is established in Mailath, Samuelson, and Shaked (1998). In their model, increasing returns to scale in the matching function drives the result. Here, we have a coordination failure.

results from separate labor markets across the observable feature. More firms then search for w workers than b workers.

Theorem 3.2 *An equilibrium exists where all w type workers draw from the high type distribution and all b type workers draw from the low distribution.*

Proof: The existence of multiple equilibria was already established by Theorem 3.1. Hence, choose c as in Theorem 3.1 and have workers of type w be in the 'good' equilibrium and have workers of type b be in the 'bad' equilibrium. QED.

Since fewer firms are searching, the unemployment rates must be higher in the bad equilibrium. Further, wage inequality is higher in the good equilibrium because the variance of θ is higher in the high distribution. Wage inequality also exists across the exogenous characteristic, with the wage inequality increasing when old.

I now examine how wage inequality across groups is affected when an equal opportunity law is put into place. The equal opportunity law is given by Definition 3.1.

Definition 3.1 *An equal opportunity law forces the labor markets for the two types of workers to be pooled. That is, firms are no longer allowed to search for workers by type.*

With an equal opportunity law of the type described above, we now have the following theorem:

Theorem 3.3 *Under the equal opportunity law, there will be no wage inequality across the observable feature among the entering cohort of young workers or any cohort thereafter.*

Proof: The two groups of entering workers are identical. Since they face the same labor market, they will behave identically, each having the same expectation on the amount and types of training and on expected wages. QED.

Note, however, whether the economy converges to the investment equilibrium or the no investment equilibrium will depned upon the set up of the problem.

Suppose the problem is set up such that the economy converges to the 'good' equilibrium.[15] In the short run, the effect of the equal opportunity law on the wages of old w types is negative.

Theorem 3.4 *Earnings will fall for old workers that were previously in the good equilibrium.*

[15] This follows directly from the proof of Theorem 3.1, where c is now set such that b types are just indifferent between investing and not investing and choose to not invest. All w type workers invest.

By pooling the labor markets we are adding a greater proportion of workers who were unemployed when young than what currently existed in the good equilibrium. Firms have lower expected profits on this group of workers. Hence, changing the proportion of workers towards one in which there are more old workers who were unemployed means means that there must now be less firms per worker. But this means that the probability of matching with a firm also falls. Since the present value of earnings is increasing in the probability of matching, workers who were previously in the steady state of the good equilibrium see there earnings fall in the short run.

4 State Persistence of Investment Costs

In the previous section I showed that, with an equal opportunity law, wage inequality would disappear across the observable feature. I now outline the effects of an equal opportunity law when the investment decisions of the parents affect the investment decisions of the children. I show that, in the short run, wage inequality will still exist across the observable feature and that those workers who were previously in the 'good' equilibrium are hurt in the short run by the equal opportunity law even if the economy eventually converges to the good equilibrium.

In particular, let there now be two costs c_H and c_L to making the investment. Workers with costs c_H never invest, while workers with costs c_L invest only in the good equilibrium. Now, the I young workers are the children of the I old workers and the draw on c for the young workers depends upon the decisions of their parents. In particular, the probabilities of drawing a particular value of c are given by:

	Child's Investment Costs	
	c_H	c_L
No Investment by Parent	$1-\pi_1$	π_1
Investment by Parent	π_2	$1-\pi_2$

where π_1 and π_2 are both less than a half. This is basically a 'role models' argument: children who see their parents invest find it less costly to invest themselves. With separate labor markets and b and w type workers in the bad and good equilibrium, respectively, the proportion of the two types who have the low costs are π_1 and $\pi_1/(\pi_1+\pi_2)$.[16] Clearly, this leads to more w type workers having low costs of investment.

[16] That an equilibrium of this type exists follows directly from the proof of Theorem3.1.

I now consider the effect of an equal opportunity law (of the same type described in the previous section) on wage inequality across groups. I assume that parents do not care about the earnings of their children. I further assume that all workers invest if their cost is low and do not invest if their cost is high. That is, I assume that the economy will eventually converge to the 'good' equilibrium should an equal opportunity law be implemented.[17]

Theorem 4.1 *Under the conditions described above, an equal opportunity law will not remove wage inequality for the next generation. Wage inequality across the types falls over time.*

Proof: The b types were previously in the bad equilibrium and therefore would never invest. This implies that the next generation of b type workers will have a greater proportion with the high cost of investment than their w type counterparts and therefore less b types will invest. As time passes, the distribution of costs for the b types converges to that to the stationary distribution for the w types. However, in the short run, wage inequality across the observable feature remains as those who make the investment make more than those who do not. As the b types distribution of investment costs converges to that of the w types so too do the average wages of b types converge to the average wages of w types. QED.

As in the previous section, the effect of the equal opportunity law on the earnings of w types is negative. Less firms are willing to enter because the expected value of θ has fallen.[18] As the b types converge to the stationary distribution, the present value of earnings for the w types increase as well due to firm entry.

5 Affirmative Action

Since wage inequality remains across the types and w types are hurt by the b types lack of investment, there may be programs one would wish to implement which accelerate the convergence to the good equilibrium.[19] In particular, w types may be supportive of a program which subsidizes the investment of some b types. The reason for this is that higher investment by b types means more firms will enter. With more firms entering, the expected present value

[17] Again, this is possible from the proof of Theorem 3.1.

[18] This is in addition to the fall in earnings due to a higher percent of the workforce being unemployed when young.

[19] 'Convergence' implies that wage inequality across the observable feature is eliminated.

of earnings for w types increases. I show that w types may want to contribute to this program initially, but in each subsequent generation the amount of the contribution will fall with a possible outcome that w types would like to stop contributing to the program before wage inequality across the types is removed.

In order to analyze the effect of the program, I will assume that the person who decides on the amount of the subsidy is young and is of type w.[20] Everyone must contribute an equal amount a where a is set by the decision maker. The total subsidy is then divided in such a way that a set number of b type workers receive just enough of a transfer such that if they were of the high cost type they would just be indifferent between making an investment. Cost types are unobserved, implying that those b types who also have low costs of investment do not change their decisions given a subsidy. Hence, the subsidy is only effective at raising $E(\theta)$ for a portion of the individuals who received it.

Theorem 5.1 *It is possible to set the parameters of the model such that the decision maker will choose to subsidize the investment of some b types.*

A key feature of the model is that the higher the percentage of b workers who have high costs, the more willing w type workers are to subsidize the investment decisions of b types, all else equal. This can be seen from examining how changing a affects $E(\theta)$.

$$\frac{\partial E(\theta)}{\partial a} = \frac{p_{bH}(\theta_H - \theta_L)}{c_H - c^*} \tag{14}$$

where p_{bH} is the percent of young b type workers who have the high cost of investment. Hence, a greater percentage of the subsidy is wasted from the w types perspective the higher the percentage of b types workers with c_L. This is why agreeing to a subsidy for w types, which also has positive effects for all workers, would be more difficult as a higher portion of them would make the investment anyway. What this implies for b types is that the percentage of b types receiving the subsidy will be falling over time as more and more b types would make the investment without the subsidy. Eventually the optimal amount number of recipients could then fall to zero and this may occur before earnings inequality across the two types is eliminated.

[20] This could be a median voter or the median voter of the young w types if, say, only the young w types were paying the subsidy.

6 Observables Investments

One of the unattractive features of this vein of the statistical discrimination literature is determining what constitutes the investment. Is it schooling? But schooling is an observable. Generally speaking, we can divide investments into two classes: observable (such as education) and unobservable (such as human capital accumulated due to 'effort'). The difference between the two for this paper is that firms are able to search over observable investments. That is, there is a separate labor market for those who make observable investmetnts. Here, I show how to incorporate investment in observables into the model.

Let workers make an observable investment decision which affects their productivity. Again let θ_i' indicate the worker's initial productivity. True productivity is then defined by θ_i where $\theta_i = [1 + \gamma_o(o = 1)]\theta_i'$, $\gamma_o \in \Re_+$. $o = 1$ then refers to whether or not the worker has made the investment.[21] Further, by making this investment, workers put themselves in a different labor market. Hence, in addition to a market for w types and b types, as well as a labor market for the young and old, there would then be separate labor markets within these groups for whether or not the worker has made an observable investment.

Suppose the only decision is whether or not to make an observable investment. In this case, there is only one equilibrium in each market as one's investment costs do not depend upon the investments of others. Hence, the investment decision would be based upon whether the returns to being in the other labor market outweighed the costs of making the investment. This condition reduces to:

$$\beta E(\theta|o = 1)\left(q_1^o + \frac{1}{1+r}\left[q_1^o(1-s)(1+\gamma_e) + sq_1^o q_2^o(1+\beta\gamma_e) + (1-q_1^o)q_2^o\right]\right) \geq$$
$$E(\theta|o = 0)\left(q_1^n + \frac{1}{1+r}\left[q_1^n(1-s)(1+\gamma_e) + sq_1^n q_2^n(1+\beta\gamma_e) + (1-q_1^n)q_2^n\right]\right) + c \qquad (15)$$

The superscripts on the probabilities dictate whether the individual is (or is not) in the market with those who have made the observable investment. Here, o superscripts indicate that the individual is in the market with the investment group while n superscripts indicate that the individual is in the group that has not made the investment. Note that the probabilities of matching conditional on investing must be greater than the corresponding probabilities conditional on not investing as firms prefer workers who invest. Hence, the labor market must be tighter for workers who make an observable investment.

[21] It is important to remember that the worker's initial productivity is not known to the worker until after the investment is made.

The next step is to add the investment decision regarding the unobservables. This involves examining the probabilities of finding a match conditional on choosing or not choosing to invest in the observables. Conditional on the investment decision in the observables, the worker's decision as to whether or not to invest in the unobservables is the same as in equation (13), where the equilibrium probabilities of matching when old and young again depend upon whether the worker has invested in the observables. Hence, the cutoff investment cost is again increasing in q_1 and q_2 (see Lemma 3.1). Further, the zero profit conditions facing searching firms imply that *in every market* the higher the investment in the unobservables, the more searching firms in the market. This follows directly from Lemma 3.2.

The conditions for Theorem 3.1 are then satisfied across markets. This implies that there may be multiple equilibria in both the markets for those who made the investment in the observable and for those who did not make the investment in the observable. Hence, a possible equilibrium would be to have some b types make the observable investment but none make the unobservable investment. On the other hand, some w types may make the observable investment but all w types may make the unobservable investment.

Again, there may be multiple equilibria. In particular, it is possible to construct the model such that only low cost workers (for both observable and observable investments) invest. However, it is also possible to have an equilibrium where those who have low costs for the observable investment do make the observable investment but neither the high unobservable cost people or the low unobservable cost people make the investment.

In order to clarify further the mechanism by which b types do not make unobservable investments even if observable investments are attractive, we reconsider the case where the investments made by the parents influence the investments made by the children. Consider the following transition matrix:

	Child's investment costs			
	c_{Hu}, c_{Ho}	c_{Hu}, c_{Lo}	c_{Lu}, c_{Ho}	c_{Lu}, c_{Lo}
Parent does not invest	π_1	π_2	π_2	π_2
Parent invests in observables only	π_2	π_1	π_2	π_2
Parent invests in unobservables only	π_2	π_2	π_1	π_2
Parent invests in both	π_2	π_2	π_2	π_1

Here, $\pi_1 > \pi_2$ establishs state persistence. Further, since this a transtion matrix, we must have $1 - \pi_1 = \frac{\pi_2}{3}$.

In the stationary distribution of the good equilibrium, where low costs for the unobservable investment imply that an investment in the unobservables is made, one quarter of the people are in each state. This implies that half of the individuals make the observable investment. Further, half also make the unobservable investment. This is the equilibrium for w types. However, in the bad equilibrium, the incentives to invest are such that they will only invest in the observables. This is where b types reside. Hence, while half of b types invest in the observables, none invest in the unobservables.

This model then has many implications for the wage profiles of w types and b types. The difference between earnings of w types and b types conditional on the observable investment indicates that, provided the relative probabilities of finding a job in the two market are sufficiently similar, the gap across the types will be larger for those who made the observable investment. This is because the observable investment increases in value with an unobservable investment and it is more costly for workers to be unemployed if they have made the observable investment.

These results hold under much less stringent conditions on the transition matrix. In fact, a more reasonable transition matrix would place higher probabilities on investments happening together: low costs for one investment means that it more likely the individual has low costs for the other investment. But this means that w types who choose the observable investment are more different from their b type counterparts than those who do not choose the observable investment. This is because b types, whether they make the observable investment or not, find it optimal to not invest in the unobservable investment while w types are more likely to make the unobservable investment if they have made the observable investment.[22]

7 Empirical Predictions

The model presented here generates many empirical predictions on how experience, education, and wages are affected by statistical discrimination. In particular, consider the case

[22]Other equilibria are also possible. For example, it may be the case that w type high cost workers find it optimal to invest in unobservables as long as they are low cost workers for the observables. This distinction could also take place for b types with investment in unobservables for low cost workers provided they also low cost workers for the unobservables. The key is that, conditioning on the market, for example young workers who have made the investment in the observables, a greater percentage of w types make the investment in the unobservable.

before the 1964 Civil Rights Act as a time when statistical discrimination was present. That is, an economy which began with slavery transitioned into one with statistical discrimination. With the 1964 Civil Rights Act, the labor market for whites and blacks became pooled. If this is correct, then we would expect to see certain trends in the data. I describe these trends in the rest of this section.

7.1 Pre-1964

Whites should have steeper wage profiles than blacks, even after controlling for education. This occurs for two reasons. The first is that human capital investments are complements here. A worker who has made an investment (either observable or unobservable) receives a bigger boost from the increase in human capital due to experience than a worker who has not made an investment. Since more white workers find it attractive to make investments (both observable and unobservable), their profiles will naturally be steeper. This will hold within educational groups because of the unobservable investments. The second reason for the steeper profiles is that whites have better job markets. Hence, the on the job training they get is more general in that it is more likely to be used than if the individual was black. This is because blacks have a lower probability of getting a job. The higher the probability of getting a job in the second period, the steeper the wage profile.

Whites should have larger within group variances than blacks. This would need to hold even after controlling for education. With whites making unobservable investments, this increases the variance on their earnings. Further, the increase is even larger when old as the on the job portion of the human capital dampens relative wages for the highly productive workers when young.

Whites should have a higher rate of return to education than blacks. Blacks have a higher probability of being unemployed. This is more damaging in the educated market because their skills are actually worth more. Further, because of the positive interaction on wages between observable and unobservable investments, whites making the observable investment receive a higher return on their unobservable investment. Since blacks do not make the unobservable investments, the wage gap would be expected to be higher the higher the educational level.

7.2 Post-1964

All of the pre-1964 trends should be mitigated. The pre-1964 trends are mitigated for two reasons. First, blacks now have an incentive to make unobservable investments. This will lead to an increase in the variance of black earnings, particularly when old. Second, the better job prospects for blacks also increase the steepness of their wage profiles. The human capital they receive from on the job training becomes more general as the labor market improves.

Higher returns to education should be observed for both blacks and whites. This last statement holds only if making the observable investment is correlated with making the unobservable investment. What is happening here is that whites are hurt by the pooling of the labor markets because blacks have not made the unobservable investment. However, this will be much more relevant for workers who have not made the observable investment as there is a greater influx of black workers into this market. A natural way to test this is by comparing changes in the college premium (and wage inequality in general) across regions of the country. The pooling of the labor markets will be much more relevant in the South, where there is a much larger percentage of blacks, than, say, the Midwest. Hence, we would expect the returns to education to increase the most in the South under the assumption that workers and firms faces costs to moving which are prohibitive in the short run.

8 Conclusion

I develop a matching model where workers make an unobservable investment decision before entering the labor force. I show that multiple equilibria may result. The key feature is that firms are unable to target their search to workers who have made the unobservable investment. The 'good' equilibria is one in which the number of searching firms is high and workers make the investment to improve their productivity. The 'bad' equilibrium has a smaller number of searching firms with no workers making the investment. Statistical discrimination can result if firms are able to search on some observable characteristic which has nothing to do with the cost of investment (i.e. b types may be in the 'bad' equilibrium and w types in the 'good' equilibrium). The results are shown not to change when workers are also allowed to make an observable investment which firms are able to search on.

We can then ask how the equilibrium changes with an equal opportunity program whereby firms must search equally hard for the two groups. Wage inequality across the observable characteristic is eliminated from the entering generation on with the equal opportunity program. However, the equal opportunity program only guarantees that both groups will be in the same equilibrium, not whether the equilibrium is the good one or the bad one. Further, even if the good equilibrium is the one that is being converged to, the short run effect for the w types is lower wages over their lifetime.

An equal opportunity program does not immediately eliminate wage inequality across types if one's investment decisions depend upon the investment decisions of one's parent. That is, let the I entering workers be the children of the previous generation. Suppose now that the cost of investment for the child depends upon whether the parent made the investment. In this case, wage inequality across types continues in the short run even with firms searching equally hard for both types of workers. This is because children of workers who were previously in the bad equilibrium are less likely to make the investment because their parents did not make the investment.

Since the implementation of an equal opportunity law means less firms per worker than in the good equilibrium, w type workers may support an affirmative program where some b type workers have their investment subsidized. Subsidizing the investment of b types means that more firms will enter. Hence, the gains all workers receive from more firms entering may exceed the cost of subsidizing some of the b types investment. However, the more likely it is that b types will invest without the subsidy, the lower the optimal number of subsidy recipients from the w types perspective. Over time, the amount of the subsidy will go to zero as the distribution of b type investment costs converges to the w type distribution.

The model can be tested empirically by examining the wage profiles of blacks and whites. In particular, whites should have steeper profiles than their black counterparts (both unconditional and conditional on making the investment) as well as higher variances for wages within each cohort. Further, if statistical discrimination was removed from, say, the 1964 Civil Rights Act, the variances of black wages should be increasing over time, though always smaller than the corresponding variances for whites. Black wage profiles since the statistical discrimination was removed should be steeper. Finally, the wage gap between blacks and whites should be increasing across educational groups pre-1964 and the returns to education across all groups should be higher post-1964.

9 Appendix

Proof of Theorem 2.1: It is useful to first show that there exist equilibrium values of q_2 and q_1 and then show how these values relate to the equilibrium values of J_2 and J_1.

Pick $q_{m'}$ and k_m, where $m \in \{1,2\}$ and m' is o when m is y, etc. Note that $q_m \in [0,1]$. Define $\overline{J_m}$ as the value of J_m such that $q_m = 1$. Note further that profits in market m are decreasing in J_m, all else equal. Profits must then be positive at $\overline{J_m}$ for the equilibrium value of q_m to equal one. If the equilibrium value of q_m is zero, then it must be the case that expected profits are negative at $J_m = 1$. If the equilibrium value of $q_m \in (0,1)$, then it must be the case that expected profits are zero.

Suppose for a given q_2, k_1, $q_1 \in (0,1)$. Then q_1 is such that:

$$0 = \frac{E(\theta)(1-\beta)q_1}{f(q_1)} \left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2\right]\right) - k_1$$

Define g as the inverse function of $\frac{f(q_1)}{q_1}$.[23] This implies for $q_1 \in (0,1)$ that the equilibrium value of q_1 follows:

$$q_1 = g\left(\frac{E(\theta)(1-\beta)\left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2\right]\right)}{k_1}\right)$$

Define q_1^* as the equilibrium value of q_1^*. Now, conditional on any value of q_2 and k_1, q_1^* is given by:

$$q_1^* = \begin{cases} 1 & \text{if } \frac{E(\theta)(1-\beta)}{\overline{J_1}}\left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2\right]\right) - k_1 > 0 \\ 0 & \text{if } E(\theta)(1-\beta)\left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2\right]\right) - k_1 < 0 \\ g\left(\frac{E(\theta)(1-\beta)\left(1+\frac{1}{1+r}\left[(1-s)(1+\gamma_e)+s\beta\gamma_e q_2\right]\right)}{k_1}\right) & \text{otherwise} \end{cases}$$

Note that, for any positive value of k_1, q_1^* is a continuous function of q_2.

Define q_2^* as the equilibrium value of q_2 conditional on q_1 and k_2. As above, we can write

[23] I showed in section 2 that $\frac{f(q_1)}{q_1}$ is continuous and increasing in q_1 for $q_1 \in (0,1)$. Hence, by the inverse function theorem, g exists.

q_2 as a function of q_1.

$$q_2^* = \begin{cases} 1 & \text{if } \frac{E(\theta)(1-\beta)}{[q_1 s+(1-q_1)]\overline{J_2}}[sq_1(1+\gamma_e)+(1-q_1)] - k_2 > 0 \\ 0 & \text{if } \frac{E(\theta)(1-\beta)}{[q_1 s+(1-q_1)]}[sq_1(1+\gamma_e)+(1-q_1)] - k_2 < 0 \\ g\left(\frac{E(\theta)(1-\beta)[sq_1(1+\gamma_e)+(1-q_1)]}{[q_1 s+(1-q_1)]k_2}\right) & \text{otherwise} \end{cases}$$

Again, for any positive value of k_1, q_2^* is a continuous function of q_1.

Let $\phi_2(q_2) : [0,1] \to q_1^*$ and $\phi_1(q_1) : [0,1] \to q_2^*$ and be given by the displayed expressions for q_1^* and q_2^*. Since ϕ_2 and ϕ_1 both map from $[0,1]$ to $[0,1]$ and are continuous in q_2 and q_1 respectively, by Brouwer's fixed point theorem there exists an equilibrium pair $\{q_1^*, q_2^*\}$. This holds for all pairs $\{k_1, k_2\}$.

Now that we have the equilibrium probabilities of finding a match, we can use these probabilities to find the equilibrium number of firms, J_m^*. If $q_m^* = 0$, then $J_m^* = 0$. If $q_m^* \in (0,1)$, then J_m^* is given by $I_{um}^* f(q_m^*) = J_m^*$. If $q_m^* = 1$, note that as $J_m \to \infty$ expected profits go to $-k_m$. Hence, by the intermediate value theorem there exists a $J_m^*(q_{m'}^*)$ such that the expected zero profit condition holds. QED.

Proof of Lemma 3.1: Differentiating the reservation cost of investment with respect to both q_2 and q_1 yields:

$$\begin{aligned} \frac{\partial c^*}{\partial q_2} &= \frac{\beta(\theta_H - \theta_L)}{1+r}\left[sq_1(1+\beta\gamma_e) + 1 - q_1\right] \\ \frac{\partial c^*}{\partial q_1} &= \beta(\theta_H - \theta_L)\left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s(1+\beta\gamma_e)q_2 - q_2\right]\right) \end{aligned}$$

Since both q_1 and q_2 are less than one, $\frac{\partial c^*}{\partial q_m} > 0$. QED.

Proof of Lemma 3.2: Define F_1 and F_2 from the zero profit conditions:

$$\begin{aligned} F_1(q_1, q_2) &= \frac{E(\theta)(1-\beta)q_1}{f(q_1)}\left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2\right]\right) - k_1 \\ F_2(q_1, q_2) &= \frac{E(\theta)(1-\beta)q_2}{[q_1 s + (1-q_1)]f(q_2)}[sq_1(1+\gamma_e) + (1-q_1)] - k_2 \end{aligned}$$

Using the implicit function theorem, we can then write:

$$\begin{pmatrix} \frac{\partial q_2}{\partial E(\theta)} \\ \frac{\partial q_1}{\partial E(\theta)} \end{pmatrix} = -\left(\frac{\partial F_2}{\partial q_2}\frac{\partial F_q}{\partial q_1} - \frac{\partial F_1}{\partial q_2}\frac{\partial F_2}{\partial q_1}\right)^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial q_1} & -\frac{\partial F_2}{\partial q_1} \\ \frac{-\partial F_2}{\partial q_2} & \frac{\partial F_1}{\partial q_2} \end{pmatrix} \begin{pmatrix} \frac{\partial F_2}{\partial E(\theta)} \\ \frac{\partial F_1}{\partial E(\theta)} \end{pmatrix}$$

The first term in parentheses is positive by Assumption 3.1. The elements of the 2×2 matrix were calculated at the beginning of section 3 and are all negative. All that is left is to show that the derivatives of the profit functions with respect to $E(\theta)$ are positive. The derivatives are:

$$\frac{\partial F_1}{\partial E(\theta)} = \frac{(1-\beta)q_1}{f(q_1)}\left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2\right]\right) > 0$$

$$\frac{\partial F_2}{\partial E(\theta)} = \frac{(1-\beta)q_2}{[q_1 s + (1-q_1)]f(q_2)}[sq_1(1+\gamma_e) + (1-q_1)] > 0$$

QED.

Proof of Theorem 3.1: Let the number of searching firms when everybody (nobody) invests lead to a probability of matching of q_{mH} (q_{mL}) where $m \in \{1,2\}$. By Lemma 3.2, $q_{mH} > q_{mL}$ $\forall m$. I show that there is a range for c such that workers invest if the probability of matching is given by q_{mH} and do not invest if the probability of matching is given by q_{mL}. Note by Lemma 3.1 c^* is increasing in both q_1 and q_o. Hence, $c^*(q_{yL}, q_{oL}) < c^*(q_{yH}, q_{oH})$. With no heterogeneity in c and setting c to be between the two reservation investment costs, this completes the proof. QED.

Proof of Theorem 3.4: Using the firm's zero profit condition in the old market and the implicit function theorem, we can find $\frac{\partial q_2}{\partial q_1}$ and $\frac{\partial q_2}{\partial E(\theta)}$ with no feedbacks; q_1 is already set form w types being in the good equilibrium and b types being in the bad equilibrium.

$$\frac{\partial q_2}{\partial q_1} = -\frac{\frac{\partial F_1}{\partial q_1}}{\frac{\partial F_2}{\partial q_2}}$$

From the proof of Lemma 3.2, we know this is positive. Similarly,

$$\frac{\partial q_2}{\partial E(\theta)} = -\frac{\frac{\partial F_2}{\partial E(\theta)}}{\frac{\partial F_2}{\partial q_2}}$$

is also positive from the proof of Lemma 3.2. Since both q_1 and $E(\theta)$ fall from the perspective of old w types, earnings for this group falls. QED.

Proof of Theorem 5.1: The decision maker faces the following optimization problem where θ_d is his expected productivity.

$$\max_a \beta\theta_d\left(q_1 + \tfrac{1}{1+r}\left[q_1(1-s)(1+\gamma_e) + sq_1q_2(1+\beta\gamma_e) + (1-q_1)q_2\right]\right) - a$$
$$s.t. \qquad 0 \le a \le p_b(c_H - c^*)$$

Here, $c_H - c^*$ is the cost of subsidizing one individual's investment and p_b is the proportion of the population who are young and of the b type. The right hand side of the constraint restricts the maximum number of people receiving the investment to be the number of entering b types in the economy. The percent of the people in the econmy being subsidized is then given by $\frac{a}{c_H - c^*}$.

The effect a a has on profits comes through the change in $E(\theta)$. In particular, we have:

$$E(\theta) = E(\theta^0) + \frac{(\theta_H - \theta_L)(1-\pi_1)a}{c_H - c^*} \tag{16}$$

where $E(\theta^0)$ is the expectation on θ without any subsidy, $(1-\pi_1)$ is the percentage of b types who have high investment costs, and $c_H - c^*$ is the amount needed to subsidize one b type's investment.

Subsituting this expression into the zero profit condition on entry in the young market, we can solve out for a as a function of q_2 and q_1.

$$a = \frac{c_H - c^*}{(\theta_H - \theta_L)(1-\pi_1)} \left[\frac{k_1 f(q_1)}{q_1 \left(1 + \frac{1}{1+r}[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2]\right)} - E(\theta^0) \right] = g(q_1, q_2)$$

Defining q_2 as an implicit function of q_1 and subsituting in for a with the above equation, we then have a maximization problem with respect to q_1.

$$\max_{q_1} \beta\theta_d \left(q_1 + \tfrac{1}{1+r}\left[q_1(1-s)(1+\gamma_e) + sq_1q_2(q_1)(1+\beta\gamma_e) + (1-q_1)q_2(q_1) \right]\right) - a \tag{17}$$
$$s.t. \qquad 0 \le g(q_1, q_2(q_1)) \le p_b(c_H - c^*)$$

By differentiating equation (17) with respect to q_1 and evaluating at q_1^0, the probability of finding a match when young without the subsidy, we can see whether parameter values exist such that this number is positive. If there is, then our proof is complete. Therefore, I want to show that parameter values exist such that:

$$\beta\theta_d \left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s(1+\beta\gamma_e)\left(q_2(q_1) + \frac{\partial q_2}{\partial q_1} q_1 \right) - q_2(q_1) + \frac{\partial q_2}{\partial q_1}(1-q_1) \right]\right)$$
$$- \frac{c_H - c^*}{(\theta_H - \theta_L)(1-\pi_1)} \left[\frac{k_1 \left(\frac{\partial f(q_1)}{\partial q_1} q_1 - f(q_1) - \frac{q_1 s\beta\gamma_e \frac{\partial q_2}{\partial q_1}}{1 + \frac{1}{1+r}[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2(q_1)]} \right)}{q_1^2 \left(1 + \frac{1}{1+r}[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2] \right)} \right] > 0$$

when the function is evaluated at q_1^0

Note that all of the terms involving $\frac{\partial q_2}{\partial q_1}$ are positive. Hence, a sufficient condition for the above inequality to hold is given by:

$$\beta\theta_d \left(1 + \frac{1}{1+r}\left[(1-s)(1+\gamma_e) + s(1+\beta\gamma_e)q_2(q_1) - q_2(q_1)) \right] \right)$$

$$-\frac{c_H - c^*}{(\theta_H - \theta_L)(1-\pi_1)} \left[\frac{k_1 \left(\frac{\partial f(q_1)}{\partial q_1} q_1 - f(q_1) \right)}{q_1^2 \left(1 + \frac{1}{1+r}[(1-s)(1+\gamma_e) + s\beta\gamma_e q_2] \right)} \right] > 0$$

Note that the first term is positive while the second term is negative as $\frac{\partial f(q_1)}{\partial q_1} q_1 - f(q_1)$ is positive from the proof of Lemma A.1. However, the second term can be made arbitrarily small by setting c_H to be ϵ greater than c^*, where c^* is defined as the reservation cost conditional on all low cost workers investing regardless of the value of q_1^0. QED.

References

[1] **Aigner, Dennis J. and Glen C. Cain.** "Statistical Theories of Discrimination in the Labor Market." *Industrial and Labor Relations Review*, (1977).

[2] **Arrow, Kenneth J.** "The Theory of Discrimination" in *Discrimination in Labor Markets,* Orley Ashenfelter and Albert Rees, editors. Princeton, NJ: Princeton University Press 1973.

[3] **Antonovics, Kate.** "Statistical Discrimination and Intergenerational Income Mobility." manuscript, 1999.

[4] **Arcidiacono, Peter.** "Firm Specific and General Training in a Matching Model." manuscript, 1999.

[5] **Becker, Gary.** *The Economics of Discrimination*, University of Chicago Press, 1957.

[6] **Chung, Kim-Sau.** "Affirmative Action as an Implementation Problem." manuscript, 1998.

[7] **Coate, Stephen and Glen Loury.** "Will Affirmative Action Eliminate Negative Stereotypes?" *American Economic Review* (1993).

[8] **Mailath, George; Samuelson, Larry and Avner Shaked.** "Endogenous Inequality in Integrated Labor Markets with Two Sided Search." manuscript, 1998.

[9] **Lundberg, Shelly and Richard Startz.** "Private Discrimination and Social Intervention in Competitive Labor Markets." *American Economic Review* 73 (3:1983).

[10] **Moro, Andrea and Peter Norman.** "Affirmative Action in a Competitive Economy." manuscript, 1999.

[11] **Phelps, Edmund.** "The Statistical Theory of Racism and Sexism." *American Economic Review* (1972).

[12] **Pissarides, Christopher.** "Loss of Skill During Unemployment and the Persistence of Employment Shocks." *Quarterly Journal of Economics* (1992).