Intergenerational Externality and the Persistence of Affirmative Action

Debipriya Chatterjee*
Brown University

Abstract

This paper examines affirmative action policies in a dynamic setting. In a world where parent's education reduces the child's cost of exerting effort, I derive and dicuss the optimal affirmative action policy. The main finding of the paper is a characterization of conditions under which affirmative action leads the economy to a "patronization trap", a scenario in which such policies continue to be persistent as eliminating them would involve incurring significant social loss over a prolonged time.

1 Introduction

Some important policy questions with regard to affirmative action pertain to: how much and how long? What should be the magnitude of the intervention policy and how should the policy behave over time?

This paper attempts to analyze affirmative action from a dynamic perspective. In a world characterized by spillovers of parental education (or qualification) on the child's cost of exerting effort (or ability), how should positive discrimination policies behave over time? I consider a general

*Email: Debipriya underscore Chatterjee at brown dot edu. I am deeply indebted to my advisor Prof. Glenn C. Loury for his unwavering support and encouragement. I am also thankful to Prof. Roberto Serrano for providing constant guidance and ever-helpful advice. I am grateful to Prof. Pedro dal Bo for his meticulous feedback and critique. I wish to thank my colleagues and seminar participants at Theory and Race and Inequality Workshops at Brown University.

equilibrium model with imperfect information [1] and identity-sightedness [2] to study this problem. As a legacy of past discrimination, a certain group (race, caste, class) of people face a higher cost of exerting effort relative to another (privileged) group. This translates to the two groups being unequally represented in different occupations. Typically, the privileged group is a majority in the high-paying, high-status jobs and the discriminated group is largely confined in the low-paying jobs that require minimal qualification. So even in a modern society that is not characterized by discrimination, inequality in terms of occupational success and the associated improved standard of living, might persist for a very long time.

I posit that the government is concerned by the disparity across the two groups in terms of their occupational success. It believes that if such significant differences continue to persist for a very long time, then there is a possibility that it might lead to a social unrest. I do not model the forces and factors that might actually be behind such a possibility in this paper. I propose that if such an episode of civil unrest does take place, then the society would incur significant loss in terms of life and property, say, and the government would like to avoid such an incident. The probability with which such an incident might happen is determined by the difference across groups in their representation in the high-level jobs of the economy.

The main result of the paper states that, under certain conditions, affirmative action might serve to discourage effort among the target group of workers (the group whose representation is of concern), which goes on to adversely affect the cost distributions of the future generations. Since the future generations find it even harder to exert effort, the representation of the two groups that would obtain without any affirmative action, would be even more unequal, thereby requiring the govt. to engage in more severe affirmative action. This perverse cycle can continue to persist for long and this would lead the economy to a long run equilibrium characterized by non-vanishing affirmative action policies and a low long run rate of human capital investment. Borrowing from Coate and Loury (1993)[4], I call this possibility a "patronization trap": a system in which prior affirmative action has affected incentives so adversely, that Laissez-faire would lead to a very unequal outcome in terms of representation and expose the society to the possibility of a civil unrest. It is observed that this obtains only when the govt. launches affirmative action in a system that is

[1] Although in this set-up the decentralized economy does not attain the first best, this is not posited as a ground for government intervention. The less than first best behavior of the agents does contribute to prolonging the time it takes for convergence in the rates of investment in human capital, but intervention is motivated by unequal success at occupational front and not the difference in the rates per se.

[2] This says that group-identity information can be used in making hiring and wage payment decisions. In other words, identity is visible and contractible.

“not very” discriminatory against the target group workers to begin with.

Through the exercises in this paper, I try to shed light on the policy debate: “gradualism” versus “one-shot” intervention. I show that over-ambitious policies, that aim to erase the unequal histories of the two groups, might backfire in the sense discussed above. In such scenarios, the government should bear the social loss of unequal representations for prolonged periods and should not intervene. This would be an example of extreme gradualism, where there is practically no intervention and market forces continue to rule. Convergence in the rates of human capital investment and representation in the high-paying tasks might take long but would take place eventually. As opposed to this, are the scenarios where policy intervention works in exactly the desired manner. Affirmative action serves to encourage incentives and the positive externality from that gets passed on to the future generations who now face a better cost distribution and are more likely to exert effort. This reduces the disparity that would obtain across groups under Laissez-faire and consequently, reduces the magnitude of affirmative action that is required.

In this paper, the government behaves like a Stackelberg leader who imposes affirmative action policy on the employers. The latter then anticipate worker behavior and devise hiring policies such that the representation concerns are satisfied. The workers observe the hiring policies, deduce incentives and make effort decisions. So the govt. in choosing its policy internalizes the ensuing equilibrium and chooses the policy that maximizes its temporal objective function. The choice of a policy would typically be derived by equating the marginal inter-temporal benefit to the cost. For instance, the policy might involve sacrificing output today for a lower cost of effort tomorrow and a lower risk of the social unrest today.

Since affirmative action is practised in many different countries of the world in many different forms, a theoretical analysis of its dynamic aspects is important. For instance, the positive discrimination privileges set aside for the lower castes and backward classes in the Constitution of India were meant to expire 20 years after Independence. However, the privileges have not only not expired in the 62 years, they have actually expanded in terms of the demographics they apply to as well as in terms of the sectors that now face the mandate of providing these privileges. Initial provisions included- reservation of parliamentary seats and preferential treatment in public services, public educational institutions.[6]. Currently such preferential policies towards the backward castes and classes have to be implemented by all government supported autonomous bodies (in the employment as well as in the educational sector). It might be useful to analyze if initial provisions actually went on to breed conditions that made the case for increasing them to the current proportions.

That quota-like affirmative action policies might affect incentives adversely has been well studied in static partial [4] and general equilibrium models [9]. I propose to extend the existing insights by evaluating the long run effects of affirmative action in a world where there are positive intertemporal externalities of acquiring education. [7] considers a static problem of affirmative action in a model of taste discrimination [2] and obtains conditions under which patronization can result. It then suggests

> "*Rather than settling immediately upon the proportional target..., the regulator could instead operate a gradual policy, with the target being ratcheted up in a series of steps*".

This paper attempts to formalize and verify this conjecture in an explicitly dynamic model characterized not by taste but by statistical discrimination. [1]

The rest of the paper is organized as follows: Section 2 describes the basic model, discusses the behavior of the economic agents and establishes that the decentralized economy does not attain the first best; Section 3 introduces affirmative action, discussing the static and the dynamic aspects; Section 4 considers the possibility when affirmative action might lead to patronization trap. Section 5 discusses the results and makes concluding remarks. The proofs of all lemmas and propositions along with some details that have been left out of the text for expositional clarity, are included in the Appendix.

2 Model

2.1 Set-up

This section introduces the basic framework. There is a continuum of workers of population size normalized to unity. Each worker can be identified to belong to any one of two distinct groups, a or b. Let $\lambda^j \in (0,1)$ be the size of group j for $j = a, b$ which does not change over time. Each worker lives for one period during which she might exert effort (invest) to acquire education and get a job. At the end of the period she is replaced by a member of the next generation. Time is discrete and is denoted by $t = 0, \ldots, \infty$. Effort choice is binary, taking value 1 if the worker did invest and become qualified and 0 otherwise. A worker, of either group, has a cost of exerting effort, k, which is drawn randomly from a distribution $G_{e_{t-1}}(\cdot) : [\underline{c}, \overline{c}] \rightarrow [0, 1]$, where $e_{t-1} \in 0, 1$ denotes the qualification of the worker's parent. If the parent did invest in acquiring education, $e_{t-1} = 1$, then it is assumed that the worker *is more likely to have lower cost of exerting effort in acquiring education* as opposed to the case in which the parent did not invest in acquiring education, $e_{t-1} = 0$. Formally,

Assumption. $G_{(e_{t-1}=1)}(\cdot)$ *first order stochastic dominates* $G_{(e_{t-1}=0)}(\cdot)$.

Each worker observes his cost and makes an effort decision. Firms cannot observe workers' costs or their effort decisions. They can only observe a test performance which is a noisy signal of the worker's effort choice to base their hiring decisions. The following subsections characterize the testing technology, the behavior of the firms and the workers respectively. These subsections describe behavior of agents in any given generation which makes the time subscript, t, redundant and it shall be suppressed but will be re-introduced in the later sections when we study the equilibrium and the dynamics. As has already been mentioned in the Introduction, this paper will consider an *identity-sighted* labor market, meaning a regime in which hiring rules and wage payments can be made contingent on the worker's identity.[3]

2.2 Signal Generation

Employers cannot observe a worker's qualification and can only observe the outcome of an imperfect test. The test score θ is drawn from a distribution that depends upon whether the worker exerted effort or not. If the worker invested in education θ follows $U[\theta_1, 1]$ and if the worker did not, then θ follows $U[0, \theta_0]$ where $\theta_1 < \theta_0$. [4] Figure 1 makes it clear that the test generates essentially three outcomes:(1) "pass" when $\theta > \theta_0$, (2) "fail" when $\theta < \theta_1$ and (3) "unclear" when $\theta \in (\theta_1, \theta_0)$. Let $p_1 = Prob\{\text{"unclear"}|e = 1\}$ and $p_0 = Prob\{\text{"unclear"}|e = 0\}$. Then,

$$\begin{aligned} p_1 &= \frac{\theta_0 - \theta_1}{1 - \theta_1} \\ p_0 &= \frac{\theta_0 - \theta_1}{\theta_0}. \end{aligned}$$

Employers have *prior* beliefs about the fraction of workers in each group that choose to exert effort and become qualified. Let π^a and π^b denote the priors for groups a and b respectively. Then, the

[3]In Section 3.1 we briefly explain the consequences of assuming an identity-blind regime

[4]This testing technology is quite standard in the statistical discrimination literature. See, for example, [4], [3]

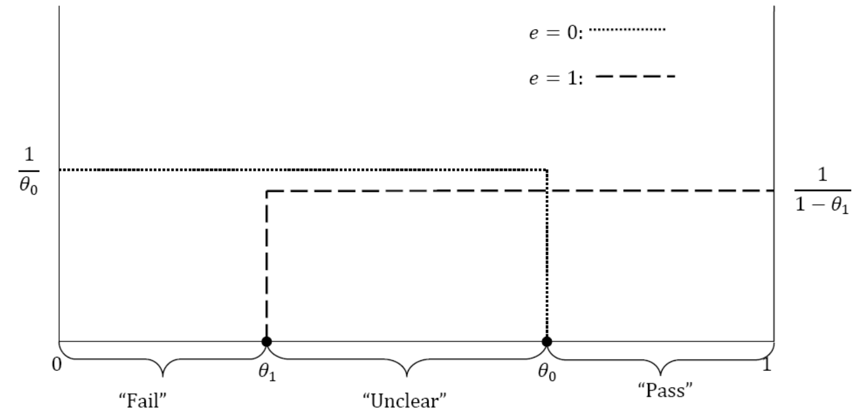


Figure 1: Signal Generation

posterior probability that a j worker with an "unclear" test result is actually qualified is:

$$\xi(\pi^j, \text{“unclear”}) \;=\; \frac{\pi^j p_1}{\pi^j p_1 + (1-\pi^j)p_0} \tag{1}$$

Without losing much generality but simplifying the algebraic calculations to a great extent, we assume that a worker who has invested in effort is equally likely to get an "unclear" on the test as a worker who did not exert effort. [5]

Assumption. $p_1 = p_0 = p$.

An immediate corollary of the above assumption is that the posterior belief about a worker's qualification whose signal is "unclear" equals the prior belief.

Corollary 1. *Given Assumption 2,* $\xi(\pi^j, \text{“unclear”}) = \pi^j$.

2.3 Employer Behavior

There is a perfectly competitive number of firms in the economy with identical production functions that use labor in two kinds of tasks, Simple and Complex, to produce output via a neoclassical production function. Only qualified workers, or workers who have exerted effort, can perform the Complex task. All workers, irrespective of their qualification, can perform the Simple task equally well. Since this is an identity-sighted labor market, firms observe a worker's group-identity and his test result and use their prior beliefs regarding the proportion of investors in that

[5] All results derived in the paper will hold for the more general case when $p_1 \neq p_0$.

group to assign him either to the Complex task or to the Simple task.

The amount of labor assigned to the Complex task is denoted by C and that assigned to the Simple task is denoted by S. However, of all the workers allotted to the Complex task, only those who are qualified count as input. Let C_q denote the *effective* labor input in the Complex task. The production technology is assumed to be Cobb-Douglas where the share of the Complex task in the output is $\beta \in (0,1)$. It should be noted that any production function that satisfies A1-A3 as listed in [10] will suffice. Thus,

$$Q \;=\; C_q^{\beta} S^{1-\beta}$$

Before analyzing the behavior of the employers observe that if labor was perfectly substitutable between the two kinds of tasks, the labor supply constraint would be linear with a slope of -1. This is indeed the case when the employers are shuffling workers with "pass" results across the tasks as those workers are qualified with certainty to perform in both the tasks. However, an "unclear" worker, belonging to group j, say, is believed to be qualified with probability $\xi^j = \pi^j < 1$. So the labor supply frontier facing the employers has linear stretches with slopes reflecting the proportion of qualified individuals in each group. The hiring policy that the employers choose would be determined by the interaction between these slopes and the marginal rate of substitution across tasks.

The problem of the employers can be described as making hiring decisions from a population of essentially four types of workers, classified as per their groups and test results: (i) workers with "pass" result, (ii) group a workers with "unclear" result, (iii) group b workers with "unclear" result and (iv) workers with "fail" result. [6] If ψ, α^a, α^b and ζ denote the probability with which workers with test result "pass", "unclear-a", "unclear-b" or "fail" are hired for the Complex task, then output is

$$Q = [\sum_j \lambda_j \left(\pi^j((1-p)\psi + p\alpha^j)\right)]^{\beta}[\sum_j \lambda^j((1-\pi^j)(1-p)(1-\zeta)+(1-\alpha^j)p+(1-\psi^j)\pi^j(1-p))]^{1-\beta} \tag{2}$$

The problem of employers is represented in Figure 2. [7] The labor supply constraint that the employers face is given by PLDCP'.

[6] Grouping this way is logical because "pass" workers of both groups are equally qualified with certainty to perform both the tasks and "fail" workers of both groups are equally unqualified to perform the Complex task for sure. The difference arises when considering the "unclear" workers in the two groups, as they happen to be qualified with different probabilities.

[7] The total labor input to the Complex task is $C = \sum_j \lambda_j \left(\pi^j((1-p)\psi + p\alpha^j) + (1-\pi^j)(p\alpha^j + (1-p)\zeta)\right)$ of which only the qualified count as input and feature as the first term in the production function.

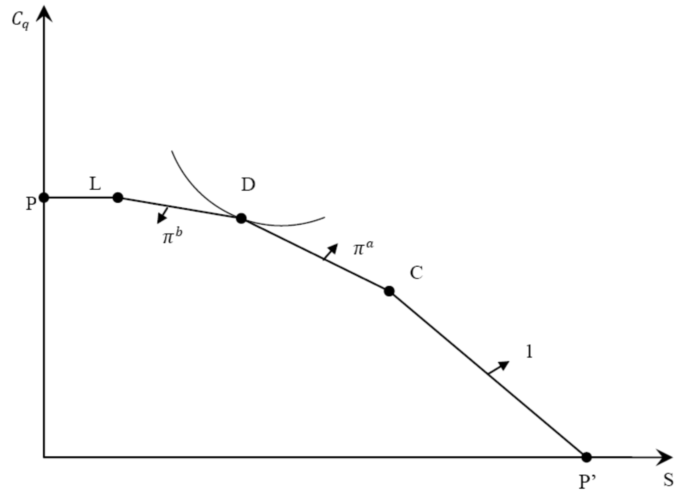


Figure 2: Labor supply frontier

Given competitive conditions and profit maximizing employers, equilibrium wages must exhaust aggregate output. So the problem of the employers reduces to choosing $\zeta, \alpha^a, \alpha^b, \psi \in [0,1]^4$ to maximize aggregate output. These choice variables pertain to the probability with which distinct categories of workers are assigned to the Complex task. Since these categories can be easily ranked in terms of the posterior productiveness of the workers in the Complex task, it should be possible to impose some structure on these choice variables without solving the employers' problem.

Lemma 2.1. *In any profit-maximizing equilibrium,*

(i) $\zeta > 0$ *only if* α^a, α^b *and* ψ *are each equal to 1.*

(ii) α^a *or* $\alpha^b > 0$ *only if* $\psi = 1$.

(iii) If employers believe $\pi^a > \pi^b$*, then in the ensuing equilibrium,* $\alpha^b > 0$ *only if* $\alpha^a = 1$.

(iv) If $\psi < 1$*, then* α^j *must equal 0, for* $j = a, b$.

The underlying intuition for these lemmas is very simple. The first one says that employers will consider hiring workers with "fail" result for the Complex task only after they have allotted *all* workers with better results, "pass" or "unclear", to the Complex task already. In the same spirit, the second lemma says that employers hire "unclear" workers for the Complex task with positive probability only after they have exhausted the supply of "pass" workers. Third states that as long as employers begin by believing that there are more qualified a workers on average than b workers, they will not hire any b "unclear" worker until they have assigned all a "unclear" workers to the Complex task. The final statement says that if the employers are hiring "pass" workers for the Complex task with a probability less than 1, then it must be the case that none of the "unclear" workers are being assigned. These results follow from the premise that a "pass" worker is certainly productive in the Complex task, but an "unclear" worker is productive only with a probability $\pi^j < 1$. The last possibility would arise if the isoquant is tangent to the constraint set on the stretch P'C. Mathematically, this happens when $\pi(1-p) > \beta$. Here the MRTS is unity because the firms are re-allocating only the "pass" workers from the Simple to the Complex task and these workers are always productive in each task with certainty. So the output maximizing allocation will be $(C_q = \beta, S = 1-\beta)$. However, this will never obtain in equilibrium because in this case the wages in the two sectors are equal and there are no returns to investing effort and acquiring education. And since investing effort is costly, no worker will have any incentive to invest, reducing π to $0 < \beta/(1-p)$ violating the beliefs of the firms. So any equilibrium in which $\psi < 1$ can be ruled out. These observations leave five possibilities for the employer's optimum:

C **Conservative** Employers operate at point C, meaning they assign only the "pass" workers to the Complex task. Thus, $\psi = 1$ and $\alpha^a = \alpha^b = \zeta = 0$. This obtains when the slope of the isoquant at C is greater than π^a, or,

$$\pi^a < \frac{1-\beta}{\beta}\frac{\pi(1-p)}{1-\pi(1-p)} \tag{3}$$

P **Partial** Employers operate on the stretch DC at a point where the slope of the isoquant is equal to π^a. They randomly assign "unclear" a workers to the Complex task with probability $\alpha^a \in (0,1)$ derived by solving

$$\frac{1-\beta}{\beta}\frac{\pi(1-p)+\lambda^a\alpha^a p\pi^a}{\lambda^a[(1-\alpha^a)p\pi^a+(1-\pi^a)(1-\alpha^a)p+(1-\pi^a)(1-p)]+\lambda^b[(1-\pi^b)+\pi^b p]} = \pi^a \tag{4}$$

Clearly, $\alpha^b = \zeta = 0$.

D **Discriminatory** Employers operate at point D. They hire all "pass" and "unclear" a workers for the Complex task but hire only "pass" b workers. This obtains when the slope of the isoquant at D is greater than π^a but less than π^b:

$$\pi^b < \frac{1-\beta}{\beta}\frac{\lambda^a\pi^a+\lambda^b\pi^b(1-p)}{\lambda^a(1-\pi^a)(1-p)+\lambda^b(1-\pi^b(1-p))} \equiv\leq \pi^a \tag{5}$$

Here $\alpha^a = 1$ and $\alpha^b = \zeta = 0$.

I **Indifferent** The employers operate at a point on the DL stretch at a point where the slope of the isoquant is equal to π^b. They hire "unclear" b workers with a probability $\alpha^b \in (0,1)$ derived by solving

$$\frac{1-\beta}{\beta}\frac{\lambda^a\pi^a + \lambda^b[\pi^b(1-p) + \pi^b p\alpha^b]}{\lambda^a(1-\pi^a)(1-p) + \lambda^b[\pi^b p(1-\alpha^b) + (1-\pi^b)(1-p) + (1-\pi^b)p(1-\alpha^b)]} = \pi^b \tag{6}$$

α^a continues to equal 1 and ζ continues to be 0.

L **Liberal** Employers operate at the point L. They are hiring all but "fail" workers for the Complex task. This obtains when the slope of the isoquant at L is less than π^b:

$$\frac{1-\beta}{\beta}\frac{\pi}{(1-\pi)(1-p)} < \pi \tag{7}$$

$\alpha^a = \alpha^b = 1$ and $\zeta = 0$.

Figure 2 depicts one of the possibilities, namely, the Discriminatory hiring regime. Since the hiring regimes are derived by comparing the slope of the isoquant to the priors held by the employers, a compact way to list regimes is:

$$(\alpha^a, \alpha^b) = \begin{cases} (0,0) & \text{if } \rho_0 > \pi^a \\ (\in(0,1), 0) & \text{if } \rho_{1,0} > \pi^a > \rho_0 \\ (1,0) & \text{if } \pi^a > \rho_{1,0} > \pi^b \\ (1, \in(0,1)) & \text{if } \rho_1 > \pi^b > \rho_{1,0} \\ (1,1) & \text{if } \pi^b > \rho_1 \end{cases} \tag{8}$$

where ρ_0, $\rho_{1,0}$ and ρ_1 are the slopes of the isoquant at points C, D and L. [8] The above list of conditions can be interpreted as partitioning the unit square formed by π^a and π^b into regions where particular hiring regimes would obtain. Since the exact nature of the partitioning would depend upon the parameters β, λ^a and p, I provide an illustration where $\beta = 0.7$, $p = 0.3$ and $\lambda^a = 0.3$. Very high values of π^a are associated with the Discriminatory or the Conservative regime, very low

[8] And are given by

$$\begin{aligned} \rho_0 &\equiv \frac{1-\beta}{\beta}\frac{\pi(1-p)}{1-\pi(1-p)} \\ \rho_{1,0} &\equiv \frac{1-\beta}{\beta}\frac{\lambda^a\pi^a + \lambda^b\pi^b(1-p)}{\lambda^a(1-\pi^a)(1-p) + \lambda^b(1-\pi^b(1-p))} \\ \rho_1 &\equiv \frac{1-\beta}{\beta}\frac{\pi}{(1-\pi)(1-p)} \end{aligned}$$

values are associated with the Liberal regime and intermediate values are associated with Partial and Indifferent regimes.

Given an analysis of the task assignment decisions, let us turn to the wage payments. w_p, w_f, and w_u^j are the wages received by a "pass" worker, a "fail" worker and an "unclear" worker who belongs to group $j = a, b$. The wage of a worker in any task must exactly equal the marginal contribution of that worker in that task. A "pass" worker in the Complex task is paid the marginal product of the Complex task, i.e. $w_p = \frac{\partial Q}{\partial C_q}$. A worker in the Simple task is always paid the marginal product of the Simple task, i.e. $w_f = \frac{\partial Q}{\partial S}$, irrespective of her test outcome. An "unclear" worker of group identity j, is productive in the Complex with posterior probability π^j; so she is paid the expected marginal product in the Complex task, i.e. $w_u^j = \pi^j \frac{\partial Q}{\partial C_q}$.

2.4 Worker Behavior

This section will analyze a worker's decision to exert effort. Workers take the hiring strategy of the firms as given and choose to exert effort and become qualified or not to exert effort and remain unqualified. For any given hiring rule, the worker calculates the net returns to exerting effort as the difference between her expected wages when she is qualified and when she is not qualified. If she is qualified, her test result would either be "pass" or "unclear", and anticipating the wages awarded for each test result, she calculates her expected wage. Similarly, if she is unqualified, her test result would be "fail" or "unclear" and she calculates her expected wage for this case. She decides to exert effort only when the difference in the expected wages is greater than her cost, k. Since the hiring strategy of the firms can be any one of the five types listed above, let us look at the workers' returns from investing in education in all the five cases:

Table 1: Returns to Investing and Not investing Effort

	e=1	e=0
Conservative	$(1-p)w_p + pw_f$	w_f
Partial	$(1-p)w_p + p\alpha^a w_u^a$, $(1-p)w_p + pw_f$	$p\alpha^a w_u^a + (1-p)w_f$, w_f
Discriminatory	$(1-p)w_p + pw_u^a$, $(1-p)w_p + pw_f$	$pw_u^a + (1-p)w_f$, w_f
Indifferent	$(1-p)w_p + pw_u^a$, $(1-p)w_p + pw_u^b$	$pw_u^a + (1-p)w_f$, $p\alpha^b w_u^b + (1-p)w_f$
Liberal	$(1-p)w_p + pw_u^a$, $(1-p)w_p + pw_u^b$	$pw_u^a + (1-p)w_f, pw_u^b + (1-p)w_f$

Given Assumption 2, the expected wage of a worker who exerts effort but gets an "unclear" test result is the same as the expected wage of a worker who does not exert effort and gets an "unclear". This simplifies the expression for returns for both groups to:

$$R(\pi^a, \pi^b) = (1-p)\{w_p - w_f\}|_{\text{at any given hiring regime}} \tag{9}$$

Whenever the employers are randomly assigning "unclear" workers of any group to either task, the expected contribution of such a worker should be same in both tasks. For instance, whenever $\alpha^a > 0$, $\pi^a \frac{\partial Q}{\partial C_q} = \frac{\partial Q}{\partial S}$ must hold. Replacing the wages by their expressions and using this last observation, returns can be expressed as:

$$R(\pi^a, \pi^b) = \begin{cases} K\rho_0^{\beta-1}(1-\rho_0) & \text{Conservative} \\ K(\pi^a)^{\beta-1}(1-\pi^a) & \text{Partial} \\ K\rho_{1,0}^{\beta-1}(1-\rho_{1,0}) & \text{Discriminatory} \\ K(\pi^b)^{\beta-1}(1-\pi^b) & \text{Indifferent} \\ K\rho_1^{\beta-1}(1-\rho_1) & \text{Liberal} \end{cases} \tag{10}$$

where $K \equiv \beta^\beta(1-\beta)^{1-\beta}(1-p)$ and ρ_0, $\rho_{1,0}$ and ρ_1 are slopes of the isoquant at points C, D and L respectively.

Lemma 2.2. *ρ_0, $\rho_{1,0}$ and ρ_1 are continuous, increasing and convex functions of* (π^a, π^b).

Detailed description of the properties of these functions and the proof of this lemma are in the appendix.

In the light of Lemma 2.2, I define the following functions:

$$\Psi(\pi^a, \pi^b) \equiv \max\{\rho_{1,0}; \min\{\pi^b, \rho_1\}\} \tag{11}$$

$$\Phi(\pi^a, \pi^b) \equiv \max\{\rho_0; \min\{\Psi(\pi^a, \pi^b); \pi^a\}\} \tag{12}$$

These functions "pick" a hiring regime by comparing the slope of the isoquant to the employers' priors. Returns can now be expressed as:

$$R(\pi^a, \pi^b) = k\Phi(\pi^a, \pi^b)^{\beta-1}(1-\Phi(\pi^a, \pi^b)) \tag{13}$$

Lemma 2.3. *$R(\pi^a, \pi^b)$ is a continuous, monotonically decreasing function of* (π^a, π^b).

Lemma 2.3 states that an improvement in the fraction of investors, of either or both groups, erodes incentives. This is an artefact of assuming the production technology to be neo-classical. An increase in the fraction of those who exert effort improves the factor ratio, thereby decreasing the marginal product in the Complex task and reducing incentives.

A worker will exert effort if and only if the returns to investment exceed cost. Since the cost of exerting effort to acquire education, k, is randomly drawn from a distribution $G^j(\cdot)$, for $j = a, b$, the proportion of group j workers that exert effort is given by $G^j(k^*) = G^j(R(\pi^a, \pi^b))$. In equilibrium, the beliefs of the employers have to be verified, or, the proportion of those who decide to exert effort from group j must equal π^j.

Proposition 1. *Given any economy $\langle p, \beta \rangle$, the priors (π^a, π^b) constitute a static Bayes-Nash equilibrium of the system if and only if*

$$\pi^a = G^a(R(\pi^a, \pi^b)) \tag{14}$$

$$\pi^b = G^b(R(\pi^a, \pi^b)) \tag{15}$$

where $R(\pi^a, \pi^b)$ is as given in (13). Such an equilibrium always exists and is unique.

Intuitively, the existence and uniqueness of the equilibrium is immediately observed since the returns function is monotonically decreasing in π^a, π^b and the cost function is non-decreasing for any group. The idea behind this kind of equilibrium is that firms expect a certain proportion of workers to exert effort and devise an assignment policy. Workers observe the hiring policy and make effort decisions. In equilibrium, the fraction that does exert effort exactly matches the expectations of the firms.

2.5 Dynamics

This section re-introduces the time subscript $t = 0, ...\infty$ and studies the behavior of the proportion of investors in the two groups over time. For notational brevity, let $R_t \equiv R(\pi_t^a, \pi_t^b)$. Assumption 1 stated that workers whose parents had exerted effort have a relatively more favorable cost distribution compared to workers whose parents had not exerted effort. The proportion of investors among those whose parents were qualified is given by $G_1(R_t)$ and the proportion of investors among those whose parents were not is given by $G_0(R_t)$. Given the qualification rates of the previous period, the fraction of investors from group j in period t is:

$$\pi_{t-1}^j G_1(R_t) + (1 - \pi_{t-1}^j) G_0(R_t)$$

which, in equilibrium, should equal π_t^j. Thus, equilibrium conditions (13) and (14) can be re-written as:

$$\pi_t^j = \pi_{t-1}^j G_1(R_t) + (1 - \pi_{t-1}^j) G_0(R_t) \quad , j = a, b \tag{16}$$

Thus the total proportion of group j workers who decide to exert effort is comprised of worker population that had qualified parents and who chose to exert effort and the worker population that had unqualified parents and who also chose to exert effort. For analytical simplicity, I assume a specific form of the cost distribution functions:

$$G_1(k) = A_1 k^2 \quad \text{and} \quad G_0(k) = A_0 k^2 \quad , A_1 > A_0$$

where $k \in [0, 1/\sqrt{A_1}]$. These cost distribution functions will be maintained for the rest of this paper. Define $A_t^j \equiv (\pi_t^j A_1 + (1 - \pi_t^j) A_0)$. Then A_t^j captures the effect of generation t faced by workers in generation $t+1$ as the convex combination of A_1, the cost parameter when the parent is qualified, and A_0, the cost parameter when the agent is not qualified, with the weight on A_1 being π_t^j. Thus the equation for the evolution of priors for any group is given by:

$$\pi_t^j = A_{t-1}^j R_t(\pi_t^a, \pi_t^b)^2, \quad j = a, b \tag{17}$$

Proposition 2. *The dynamical system in (16) has a unique stable steady state characterized by asymptotic convergence of the rates of investment in human capital for the two groups.*

The asymptotic convergence is a consequence of two assumptions- one technical and one more fundamental. The fundamental assumption is that the difference in the cost distributions faced by the two groups is a consequence solely due to the unequal histories of the two groups. In other words, I assume away any inherent differences between the groups by invoking the *anti-essentialism* axiom [9] and let the parameters A_1 and A_0 be same for the two groups. The underlying motive is to evaluate the effectiveness of affirmative action policies in undoing the legacy of the past in a world where no other factor handicaps a group or a race.[10]

The technical assumption concerns the "power" of the test which produces "unclear" result for qualified and unqualified workers with the same probability for both groups and ends up equalizing incentives across the groups in effect. So although the labor market allows for identity-sighted wage-task contracts and employers use identity information in making hiring decisions, the incentives to exert effort turn out to be the same for workers in the two groups in all hiring regimes.

[9][8], pp 5: "Anti-essentialism: The enduring and prolonged social disadvantage...is not the result of any purportedly unequal innate human capacities of the "races". Rather this disparity is a social artifact- a product of the peculiar history, culture and political economy of the ...society."

[10]In contrast to this, one can have the following hypothetical ranking of the parameters: $A_1^a = A_1^b > A_0^a > A_0^b$ as in [5] where the cost distribution is unfavorable for children of unqualified workers, the b workers being the worst off, but children of qualified workers in both groups have the same cost distribution, in the spirit of "education levels the playing field".

Since incentives are same and cost distributions differ only in terms of the fraction of workers who exerted effort in the previous period, asymptotic convergence is inevitable as the fraction of investors evolves over time.

Steady-state comparative statics yield the following observations:
(1) A high value of p gives a low value of the steady state prior. This is a reflection of the fact that a very difficult test, or a test that generates "unclear" outcomes with a high probability, eventually discourages workers from putting in effort leading to a low long run value of the fraction of investors.
(2) The higher the difference $A_1 - A_0$, the longer does asymptotic convergence take. Or, the greater the intergenerational externality effect, the longer is the persistence of unequal histories. This observation, coupled with an imminent possibility of social unrest from the disparity across groups, creates grounds for policy intervention aimed at quickening the pace of convergence. Figure 3 shows an economy where the two groups begin with very disparate initial rates of investment in education and given large intergenerational externality effects, asymptotic convergence takes very long.

[Insert Figure 3 here]

2.6 Efficiency

This section discusses the intra-generational and inter-generational inefficiencies that emerge in this decentralized economy. The intra-generational inefficiencies arise because workers while making their effort decision, consider only the effect their choice will have on the wages they receive, and ignore the effect on the probability with which they might be assigned to the Complex task, something which is determined by the qualitative nature of the "pool" of workers. Since workers do not take into account the fact that choosing to exert effort improves the fraction of qualified in the population which then gets reflected in the expected wage, they consistently under-invest. On the inter-generational front, I assume that workers are myopic. They do not recognize the effect of their investment decisions on the cost distribution of their children. This lack of foresight on the part of the parents translates to the economy being on a path of consistent under-investment. To establish this formally, we set up the planner's problem:

$$\max \sum_{t=0}^{\infty} \delta^t [Q_t - \sum_j \int_0^{\pi_t^j} (\frac{z}{A_{t-1}^j})^{1/2} \mathrm{d}z]$$

by choosing $\{\pi_t^j\}_0^\infty$, $j = a, b$.

Proposition 3. *The Laissez-faire equilibrium is characterized by less than socially optimal rates of investment.*

Proof. If $\bar{\pi}_t^a$ and $\bar{\pi}_t^b$ denote the socially optimal investment rates for groups a and b respectively, then they must satisfy:

$$\frac{\partial Q}{\partial \bar{\pi}_t^j} - \left(\frac{\pi_t^j}{A_{t-1}^j}\right)^{1/2} + \frac{A_1 - A_0}{2(A_t^j)^(3/2)} \int_0^{\bar{\pi}_{t+1}^j} \sqrt{z}\mathrm{d}z = 0 \tag{18}$$

The effect on output at time t of a change in π_t^j is:

$$\frac{\partial Q}{\partial \pi_t^j} = (1-p)\left(\frac{\partial Q}{\partial C_q} - \frac{\partial Q}{\partial S}\right) + p\alpha_t^j \frac{\partial Q}{\partial C_q} \tag{19}$$

The second term on the right of (18) captures the intra-generational inefficiency. And the last term in (17) captures the inter-generational externality, also ignored by myopic workers. Using (18) in (17) and re-arranging, the socially optimal choice of $\bar{\pi}_t^j$ should satisfy:

$$\bar{\pi}_t^j = A_{t-1}^j \left\{ \underbrace{(1-p)\left(\frac{\partial Q}{\partial C_q} - \frac{\partial Q}{\partial S}\right)}_{\text{Returns under Laissez-faire}} + p\alpha_t^j \frac{\partial Q}{\partial C_q} + \frac{A_1 - A_0}{2(A_t^j)^{3/2}} \int_0^{\bar{\pi}_{t+1}^j} \sqrt{z}\mathrm{d}z \right\}^2 \tag{20}$$

Thus, the social returns of a marginal increase in the investment rate among group j workers are more than the Laissez-faire returns as the last two terms on the right of (19) are positive. □

3 Affirmative Action

The legacy of discrimination, although no longer present in the current labor market, has caused the two groups to have very unequal histories. This is reflected in the rate at which the b and a workers are represented in the high-paying Complex task. Assuming that there is a distinct social cost to having such disparity in occupational success rates over a prolonged time, in the form of, say, a small but positive probability of social unrest, the govt. actively engages in improving the representation of the historically disadvantaged in the Complex task. However, in world where there is a positive spill-over of parental qualification, the magnitude of affirmative action that is desirable should evolve over time. For example, the children born to parents who were encouraged to exert effort and become qualified on account of affirmative action, would realize costs drawn from the better distribution and would be more likely to exert effort, thus reducing the need for affirmative action. The exercise in this subsection is to derive an optimal dynamic trajectory that the government should follow as it tries to implement its affirmative action objectives.

Definition. Representation rate for any group j in any period t, η_t^j, is defined as the fraction of its population that is assigned to the Complex task in that period.

For example, if the hiring regime is "Discriminatory", where all "pass" and "unclear" a workers and only "pass" b workers are being hired for the Complex task, then representation rates are given by:

$$\eta_t^a \;=\; \pi_t^a + p(1-\pi_t^a); \quad \eta_t^b \;=\; \pi_t^b(1-p)$$

The representation rate for a group is an increasing function of the rate of investment for that group. If $\eta_t^{j,LF}$ denote the representation rate that would obtain in equilibrium under Laissez-faire for group $j = a, b$, then, since in any period the rate of investment among group a workers is higher than that among group b workers, the latter would always be under-represented in the Complex task.

If there is a social cost to having very disparate representation rates, in the form of, say, possible social unrest, then the govt. might want to actively engage in promoting the under-represented group. This paper does not model the forces and factors behind such a social loss component. It is simply assumed that such a social unrest, if allowed to occur, would cause huge social loss in terms of life and property. I posit that, associated with any disparity in representation rates, $\Delta\eta_t \equiv \eta_t^a - \eta_t^b$, is some probability $\gamma(\Delta\eta_t) \in (0,1)$ of the social unrest. If an episode of social unrest does occur, then there is a significant loss to the society, denoted by $L > 0$. So the expected per-period loss faced by the government is $L\gamma(\Delta\eta_t)$.

Assumption. (i) $\gamma(\cdot)$ is a monotonic increasing function of the disparity in the representation rates.

(ii) $\gamma(0) = 0$.

The dynamic problem of the govt. can now be formulated as one of choosing a sequence of disparity limits, $\{\Delta\bar{\eta}_t\}_0^\infty$ that serve to maximize aggregate welfare net of the expected loss from possible social unrest.

$$\max_{\{\Delta\bar{\eta}_t\}} \sum_{t=0}^{\infty} \delta^t \left\{ (C_q^{\beta} S^{1-\beta})_t - \sum_{j=a,b} \lambda_j A_{t-1}^j \int_0^{R_t} k \, \mathrm{d}(k)^2 - L\gamma(\Delta\bar{\eta}_t) \right\} \tag{21}$$

The role of the govt. in this paper is not that of a benevolent social planner. On the other hand it is akin to that of a Stackelberg leader. The govt. imposes a sequence of disparity limit for each period and makes it mandatory for the employers to abide by it. So the employers devise their hiring strategies keeping the mandate in mind. Workers observe these hiring strategies and make effort decisions. In equilibrium, the disparity in the representation rates does not exceed the limit set by the govt. So the govt., in choosing the limit for a period, internalizes the behavior of the employers and the workers that would unfold if that limit were to be set. It then proceeds to choose the limit which maximizes it objective. Since this is a two-stage dynamic where the govt. gets to move first followed by the other economic agents, it is natural to start in the follower's world and work backwards. To this end, I first analyze the behavior of the employers and workers in a given period reacting to an exogenously given disparity limit and then come back to the dynamic problem.

3.1 Static Analysis of Affirmative Action

In any period t, employers face a disparity target, denoted by $\Delta\bar{\eta}_t$, that they cannot exceed. Clearly, this limit must be lower than the Laissez-faire disparity rate, $\Delta\eta_t^{LF}$, because the govt. is actively engaged in reducing the disparity in the representation rates of the two groups, thereby minimizing the threat of social unrest. Once again, I drop the subscript t for this section as the analysis is confined to a given period.

The problem of the employers is to maximize output subject to the affirmative action limit by choosing appropriate hiring policies. The related Lagrangian is:

$$C_q^{1/2} S^{1/2} + \mu \left[\Delta\bar{\eta} - (\eta^a - \eta^b) \right] \tag{22}$$

where $\mu > 0$ is Lagrange multiplier. Beginning from a hiring regime where none of the "unclear" b workers were being hired under Laissez-faire, the affirmative action mandate causes employers to start hiring these workers. And beginning from a Liberal hiring regime, employers must start hiring "fail" b workers. Since the implications on incentives for these two cases is very different, they each demand to be dealt separately. This section will analyze the static implications of affirmative action beginning from hiring regimes other than Liberal and section ? will discuss the remaining possibility in detail.

In a world with affirmative action (AA), let $\alpha^a \in [0,1]$ and $\alpha^b \in [0,1]$ denote the probabilities with which "unclear" a and b workers are assigned to the Complex task. Then, the resultant disparity, $(\pi^a - \pi^b)(1-p) + p(\alpha^a - \alpha^b)$, must not exceed $\Delta\bar{\eta}$. Will the disparity be less than the limit? No, because if the constraint does not bind, it is as if the employers are back in the Laissez-faire (LF) world and by construction, the disparity under LF was higher than the limit set by the govt. Since the constraint must always bind, equating the expression for disparity to the limit and solving for α^a yields:

$$\alpha^a = \alpha^b - \frac{(\pi^a - \pi^b)(1-p) - \Delta\bar{\eta}}{p} \tag{23}$$

For ease of exposition, define

$$\phi \equiv \frac{(\pi^a - \pi^b)(1-p) - \Delta\bar{\eta}}{p}$$

which captures the difference between the probabilities with which "unclear" as and bs would be hired for the Complex task. By construction, since $\Delta\bar{\eta} < (1-p)(\pi^a - \pi^b)$, ϕ is always positive. So α^a must always be less than α^b in a world with AA. Figure 3 shows this relationship between α^a and α^b. In the same square, the Laissez-faire relationship amounted to movement along the axes where α^b could be positive only if α^a had hit 1. (Lemma 2.2) [11]

So the employer's problem is now reduced maximizing $Q = C_q^{\beta} S^{1-\beta}$ by choosing an α^b in a world with AA. The labor input in the Simple task and the effective labor input in the Complex task is given by[12]:

$$C_q = \pi(1-p) + p(\pi\alpha^b - \lambda_a \pi^a \phi) \tag{24}$$

$$S = (1-\pi)(1-p) + p(1 - \alpha^b + \lambda_a \phi) \tag{25}$$

[11] In an identity-blind regime, one in which employers cannot use identity information in making hiring or wage payment decisions, the probability with which an "unclear" worker is hired must be the same for both groups. Consequently, α^a and α^b would lie on the diagonal of the unit square.

[12] Observe that: $C_q = \pi(1-p) + p(\lambda_a \pi_a \alpha_a + \lambda_b \pi_b \alpha_b)$, $S = (1-\pi)(1-p) + p(\lambda_a(1-\alpha_a) + \lambda_b(1-\alpha_b))$ and use (22).

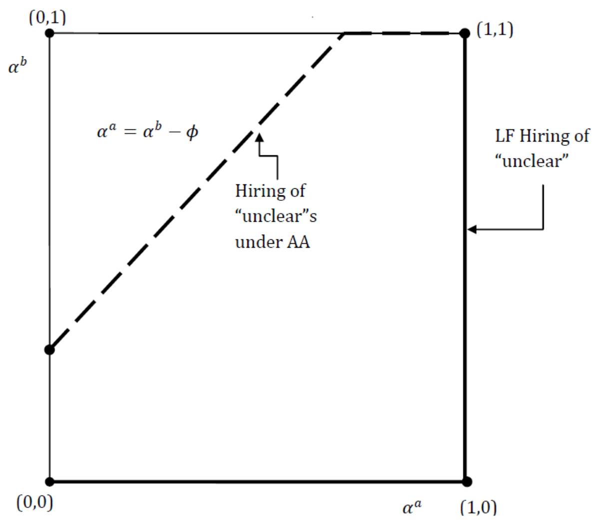


Figure 3: α^a and α^b under Affirmative Action

The first order condition is:

$$\pi = \frac{\frac{\partial Q}{\partial C_q}}{\frac{\partial Q}{\partial S}} = \frac{1-\beta}{\beta}\frac{C_q}{S} \tag{26}$$

which can be solved for α^b and using (22), one can derive α^a. To express the optimum values, I introduce α which is the probability with which an "unclear" worker is assigned to the Complex task in an *identity-blind* regime.[13] Then,

$$\alpha^a = \alpha + \frac{\lambda_a \pi^a \phi}{\pi}(1-\beta) - \phi(1-\beta\lambda_a) \tag{27}$$

$$\alpha^b = \alpha + \frac{\lambda_a \pi^a \phi}{\pi}(1-\beta) + \phi\lambda_a\beta \tag{28}$$

Observe that $\alpha^a < \alpha < \alpha^b$. To understand how the AA mandate changes the hiring decisions, let us return to the labor availability frontier that they face in Figure 4. Under Laissez-faire, the relevant frontier was PCDLP', and the employers were assigning workers to the Complex task according to the likelihood with which they were deemed to be qualified. This resulted in the five possible hiring regimes described previously. However, when faced with the AA constraint, the employers choose to abandon their preferred frontier for one which prioritizes the assignment of group b workers. To envision how such a frontier might look, imagine that the initial hiring regime is Partial. A mandate requiring employers to hire more b workers might result in the adoption of

[13] α is derived by solving $\pi = \frac{\mathrm{d}C_q}{\mathrm{d}S}$ where $C_q = \pi(1-p) + p\alpha\pi$ and $S = (1-\pi)(1-p) + p(1-\alpha)$ and is given by $\alpha = \frac{1}{p}[(1-p)(2\beta - 1 - \beta\pi) + p\beta]$.

the following scheme: an "unclear" a worker would be hired only after all the "unclear" b workers have been assigned to the Complex task. Such a frontier is depicted in Figure 4 as PCD'LP'. This would correspond to moving first along the α^b axis and then along the α^a axis in Figure 3. In theory, the employers can decide to operate at any point in the unit square defined by the α^a and α^b axes. The AA constraint, in its effort to promote the representation of b workers, forces employers to operate on the locus defined by $\alpha^a - \alpha^b = \phi$ in (22). Thus the relevant labor supply frontier is now given by PCBALP'. Note that the ray joining the points A and B has slope equal to π, so if the employers' optimum is on the AB stretch, then the slope of the isoquant must equal π as indicated in (25).

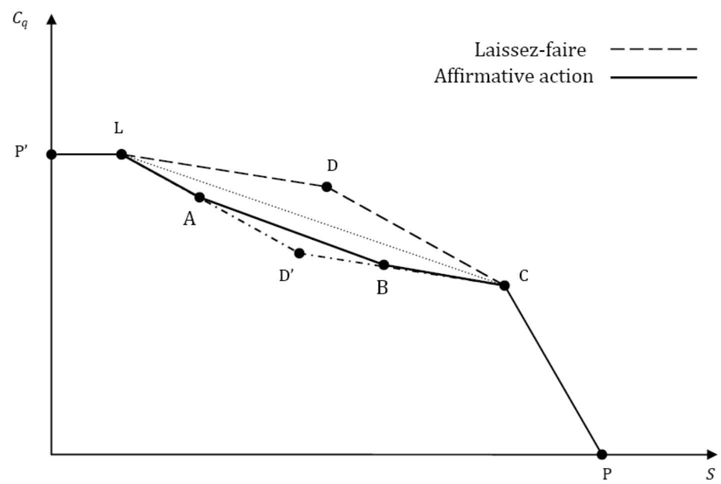
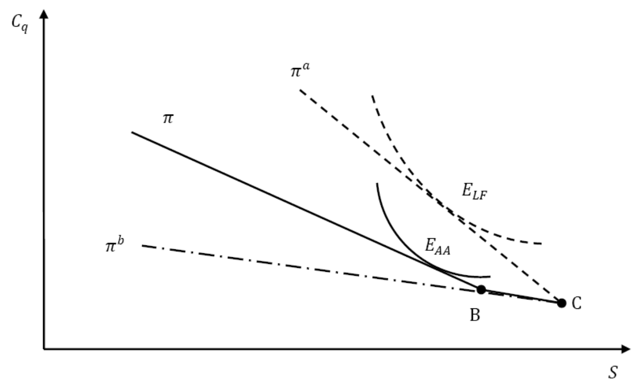


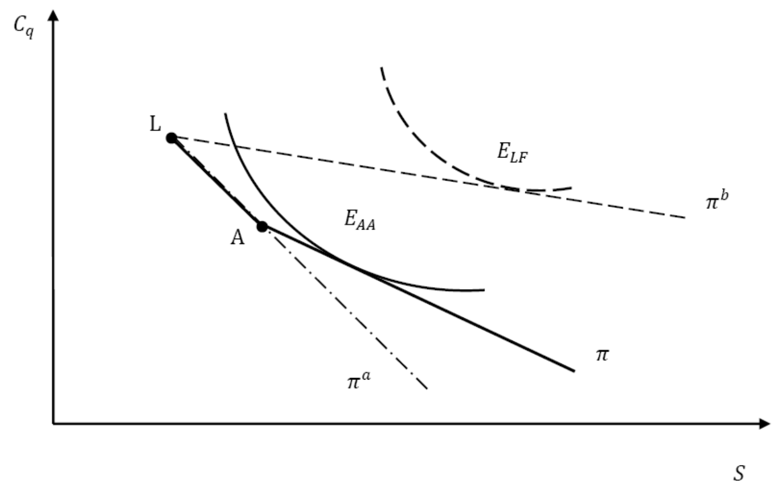
Figure 4: Labor supply frontier under Affirmative Action

Beginning from a Discriminatory or Partial hiring regime where no "unclear" b worker was being hired, the affirmative action constraint pushes the employers down to the frontier PCBALP' where the slope of the isoquant at tangency is equal to $\pi < \pi^a$, the slope of the isoquant at tangency under Laissez-faire. This would involve a lowering of the factor ratio at the optimum. Given diminishing marginal returns, this translates to an increase in returns for workers in both groups. However, beginning from an Indifferent regime where "unclear" b workers are being hired with a positive probability, the optimum under AA is at a point where the slope of the isoquant is $\pi > \pi^b$, the slope of the isoquant at tangency under Laissez-faire. This involves an increase of the factor ratio and consequently, a worsening of the returns for both types of workers. These two possibilities are depicted in Figure 5 parts (a) and (b) respectively.

To sum, in the post-affirmative action equilibrium, employers are partial towards b workers and strict towards a workers by margins that depends on the demographics, λ_a, and the magnitude of the affirmative action policy as reflected in the limit, $\Delta\bar{\eta}$. Also, the more unequal the initial regime



[part(a)]



[part(b)]

Figure 5: Employer responses to Affirmative Action

is, with respect to “unclear” b workers, the greater are the gains in representation to be had from affirmative action.

Proposition 4. *Consider all possible hiring regimes other than Liberal. When affirmative action imposes the limit $\Delta\bar{\eta}$ on employers, they react by altering their hiring policies to accept “unclear” a workers with probability given by (27) and “unclear” b workers with probability given by (28).*

Remark 1. The probability with which “unclear” a workers are hired is an increasing function of $\Delta\bar{\eta}$ and the probability with which b “unclear” workers are hired is a decreasing function of $\Delta\bar{\eta}$.

How do workers react to AA? Returns to exerting effort can always be expressed as:

$$R(\pi^a, \pi^b) \;=\; (1-p)\,\{\text{wage in Complex task} - \text{wage in Simple task}\} \tag{29}$$

Under Laissez-faire the wage paid to a "pass" worker in Complex task is the marginal product of the Complex task and that to an "unclear" worker is the marginal product adjusted for the posterior probability that the worker is qualified. Moreover, when employers are randomizing with a strictly positive probability, then the wage of an "unclear" worker in the Complex task exactly equaled his wage in the Simple task, or, the marginal product in the Simple task as shown in (4) and (6) of the hiring regime descriptions. Since the employers operated on the identity-sighted labor supply frontier PCDLP', the wage received by a j "unclear" worker assigned to the Complex task was equal to $\pi^j \frac{\partial Q}{\partial C_q}$. However, when the employers face the affirmative action constraint, they hire both a and b "unclear" workers from the population, albeit different probabilities. So the wage received by an "unclear" worker is the marginal product in the Complex task weighed by the population probability of the worker being qualified, i.e.,π, as shown in the employers' first order condition in (25). Although the probability with which the hiring decisions are made, continues to be contingent on the identity, the wage paid to an "unclear" worker is same for both the groups. Thus, replacing for the wages in (28) returns as:

$$R(\pi^a, \pi^b) \ = \ (1-p)(1-\pi)\frac{\partial Q}{\partial C_q} \ = \ K(1-\pi)\pi^{\beta-1} \tag{30}$$

The effect on returns is solely determined by the fate of the factor ratio. Reviewing the discussion above, starting from a hiring regime which did not assign "unclear" b workers to the Complex task, the factor ratio falls in the post-AA world, thereby improving incentives to invest. This is because only *some* of the b workers with "unclear" results, who were earlier working in the Simple task, are indeed qualified and contribute to the labor in the Complex task. And they have replaced a greater number of a "unclear" workers who were qualified with a higher posterior probability. This causes the labor in the Simple task to rise and the labor in the Complex task to fall, thereby reducing the factor ratio and improving the incentives. Analogously, starting from a regime where "unclear" b workers were being hired for the Complex task, the factor ratio rises in the post-affirmative action optimum. This is because the employers have to put too many workers in the Complex task, and for every unit of labor withdrawn from the Simple task, a fraction of labor is added to the Complex task, thereby raising the factor ratio and hurting incentives.

Summarizing the discussion in this section- affirmative action, of the sort considered here, causes employers to operate along a "pseudo blind" frontier where the slope of the isoquant equals π in equilibrium but the rates at which "unclear" workers from the two groups are assigned to the Complex task varies according to the required representational concerns.

3.2 Dynamics of Affirmative Action

This section discusses the evolution of the rates of investment in human capital for the two groups in a world with AA. The equilibrium priors should satisfy:

$$\pi_t^j = (\pi_{t-1}^j A_1 + (1 - \pi_{t-1}^j) A_0) R_t^2, \quad j = a, b \tag{31}$$

where R_t is as given by (29) when π is replaced by π_t. An important observation to make here is that the equation defining the equilibrium in period t is independent of the representation target chosen for period t, namely $\Delta\bar{\eta}_t$. Or, the employers anticipate certain investment rates and derive a hiring policy based on the population proportion of investors, although the exact rates at which "unclear" workers get assigned to the Complex task is different for the two groups. Equilibrium values of π_t^a and π_t^b are be derived as functions of the investment rates in the previous period, $(\pi_{t-1}^a, \pi_{t-1}^b)$ only.

As explained earlier, the government's problem, is to choose a sequence of representation targets, $\{\Delta\bar{\eta}_t\}_{t=0}^{\infty}$, in a way that maximizes aggregate social welfare net of the expected loss from a possible civil unrest. Reviewing the objective function of the govt., it is clear that a choice of representation target in period t affects social surplus in period t in the following manner: first, it affects the composition of the labor force for the two tasks and thus affects output and second, it affects the expected loss from the possible civil unrest. The government's problem can now br written as:

$$\max_{\{\Delta\bar{\eta}_t\}} \delta^t \left[Q_t - \sum_j \lambda_j A_{t-1}^j \int_0^{R_t} 2c^2 \mathrm{d}(c) - L\gamma(\Delta\bar{\eta}_t) \right] \tag{32}$$

where

$$C_q = \pi_t(1-p) + p(\lambda_a \pi_t^a \alpha_t^a + \lambda_b \pi_t^b \alpha_t^b) \tag{33}$$

$$S = (1-\pi_t)(1-p) + p(\lambda_a(1-\alpha_t^a) + \lambda_b(1-\alpha_t^b)) \tag{34}$$

$$\alpha_t^a = \alpha_t + \frac{\lambda_a \pi_t^a \phi_t}{\pi_t}(1-\beta) - \phi_t(1-\beta\lambda_a) \tag{35}$$

$$\alpha_t^b = \alpha_t + \frac{\lambda_a \pi_t^a \phi_t}{\pi_t}(1-\beta) + \phi_t \lambda_a \beta \tag{36}$$

$$\phi_t = \frac{(\pi_t^a - \pi_t^b)(1-p) - \Delta\eta_t}{p} \tag{37}$$

$$\pi_{t+1}^j = \gamma_j(\pi_t^a, \pi_t^b), \quad j = a, b \tag{38}$$

I start with considering the effect of the policy on the output. The govt. aims to reduce disparity across the two groups in the extent to which they are represented in the high-paying Complex task. So a greater desire for equality would mean a lower allowed limit, or a more "severe" mandate. It is observed that as the policy becomes severe, output falls. This effect should not come as a surprise because a lower target requires firms to hire more "unclear" b workers who are qualified with a strictly lower probability than the a workers who are being replaced in the course.

Lemma 3.1. $\frac{\partial Q}{\partial \Delta \bar{\eta}_t} > 0.$

Given this observation, the optimal choice of $\Delta\bar{\eta}_t$ for any t should satisfy:

$$\frac{\partial Q_t}{\partial \Delta \bar{\eta}_t} = L\gamma(\Delta\bar{\eta}_t) \tag{39}$$

As is clear from this condition, the choice of limit in any period does not get carried over to the next period. This is a consequence of assuming a Cobb-Douglas technology where the factor ratio continues to be determined by the population prior, π, and where AA changes the composition of the labor input in the two tasks leaving incentives unchanged.

Proposition 5. *Beginning from any initial hiring regime where "unclear" b workers are not being hired for the Complex task with probability 1, the optimal choice of $\Delta\bar{\eta}_t$ in any period t is given by the solution to:*

$$(\frac{C_q}{S})_t^{\beta-1}\left[\lambda_a(\pi_t^a - \pi_t)\left(\beta^2 + (1-\beta)^2\frac{1}{\pi_t}(\frac{C_q}{S})_t\right)\right] = L\gamma'(\Delta\bar{\eta}_t) \tag{40}$$

The optimal choice in each period should equate the marginal sacrifice of output to the marginal reduction in the probability of a civil unrest.

4 Patronization

Till now, this paper has considered the problem of setting a limit on the permitted levels of disparity in representation for hiring regimes that do not already assign all "pass" and "unclear" a and b workers to the Complex task. This section will take up the problem of the possibility: namely, how is policy making different for the Liberal hiring regime?

The Laissez-faire difference in the representation rates is now given by $\Delta\eta^{LF} = (\pi^a - \pi^b)(1 - p)$. If the govt. imposed requirement is to have a difference less than $\Delta\eta < \Delta\eta^{LF}$, then the employers will need to start hiring some “fail” b workers.[14] Let ζ denote the probability with which a b worker with “fail” result is assigned to the Complex task. Then, the labor input in the Complex task continues to equal π as before, but the labor input in the Simple task has fallen because these “fail” workers will not contribute anything to the Complex task although they were contributing valuable labor to the Simple task. The employers maximize output subject to the representation concerns by choosing appropriate hiring policies. The associated Lagrangian is:

$$\mathcal{L} = \pi_t^{\beta}((1-\pi_t)(1-p)-\lambda_b(1-\pi_t^b)(1-p)(1-\zeta_t))^{1-\beta}+\mu_t\left[\Delta\bar{\eta}_t-(\pi_t^a-\pi_t^b)(1-p)-\zeta_t(1-\pi_t^b)(1-p)\right] \tag{41}$$

where $\mu_t \geq 0$ is the multiplier. The representation constraint must bind at the optimum because the employers will not want to add any more non-productive workers to the Complex task than is required to meet the constraint. Solving for ζ from the above constraint:

$$\zeta_t = \frac{\pi^a - \pi^b}{1-\pi^b} - \frac{\Delta\bar{\eta}_t}{(1-\pi_t^b)(1-p)} \tag{42}$$

So $\mu_t > 0$ at the optimum. This can be interpreted as a “shadow tax (subsidy)” imposed on the employers for assigning an additional a (b) worker to the Complex task. The contribution of a “pass” a worker to the Complex task which equaled the marginal product in the Complex task earlier, is now reduced by μ_t, whereas that of a “pass” b worker is now increased by the same amount. Similarly, for the “unclear” a and b workers, their expected contribution in the Complex task is lowered (raised) by the multiplier. And the “fail” a workers assigned to the Simple task continue to enjoy the marginal product in the Simple task. This leaves only the “fail” b workers, some of whom are assigned to the Complex task. The physical contribution of such a worker to the Complex task is nil but since it helps the employers to meet the constraint, his economic contribution is μ_t. And since employers are randomly assigning these workers to either of the two tasks, their contributions in the two tasks must be equal. Thus, $\mu_t = \frac{\partial Q_t}{\partial S_t}$.

Given these observations about the employers’ behavior, it is now possible to evaluate the incentives faced by the workers in the two groups. The wage of an a worker placed in the Complex task is $\frac{\partial Q}{\partial C_q} - \mu_t$ and the wage of a b worker in the Complex task is $\frac{\partial Q}{\partial C_q} + \mu_t$. As discussed above,

[14] Since the “fail” workers do not contribute anything to the labor input in the Complex task for certain, it is efficient for the employers to hire just the minimal number of “fail” b workers so that they can meet the constraint.

both types of workers, if placed in the Simple task, get $\frac{\partial Q}{\partial S}$. Thus, returns are:

$$R_t^a = (1-p)(\frac{\partial Q}{\partial C_q} - \mu_t) - (1-p)\frac{\partial Q}{\partial S} \tag{43}$$

$$R_t^b = (1-p)(\frac{\partial Q}{\partial C_q} + \mu_t) - (1-p)(\zeta_t \mu_t + (1-\zeta_t)\frac{\partial Q}{\partial S}) \tag{44}$$

Thus, compared to Laissez-faire, the returns to invest have clearly decreased for a workers. For b workers, although it might appear that the policy should encourage effort, it actually decreases incentives because the increased factor ratio causes marginal product in the Complex task to fall. So beginning from a Liberal hiring regime, affirmative action aiming to lower disparity in representation causes employers to hire "fail" b workers for the Complex task which serves to reduce returns for workers in both groups.

How do the equilibrium rates of investment respond to an increase in the severity of the policy in this scenario?

Lemma 4.1. *Beginning from a Liberal hiring regime, any increase in the severity of affirmative action, that is, a lowering of the disparity limit, causes investment rates to worsen for both groups.*

So the affirmative action has served to worsen the investment behavior for both groups. Given that parental qualification spills over to affect the distribution from which the child's cost of exerting effort is drawn, it is intuitive that worsened qualification rates will materialize as worsened cost distributions for the next generation. But what happens to disparity in the representation rates in the next period? If the govt. decides to do away with affirmative action and let the system get back to a Laissez-faire equilibrium, then, is there greater or less disparity in the ensuing Laissez-faire equilibrium?

To answer this, note that the disparity in the next period can be expressed as:

$$\Delta\eta_{t+1} = (\pi_{t+1}^a + p(1-\pi_{t+1}^a)) - (\pi_{t+1}^b + p(1-\pi_{t+1}^b)) = (1-p)(\pi_{t+1}^a - \pi_{t+1}^b) \tag{45}$$

Since we are back to the Laissez-faire world in period $t+1$, so returns will now be given by $R_{t+1}^a = R_{t+1}^b = R_{t+1}$ where R_{t+1} is the expression for returns under the Liberal hiring regime. Using the equations that describe equilibrium priors, we can re-write (51) as:

$$\Delta\eta_{t+1} = (1-p)(A_1 - A_0)(\pi_t^a - \pi_t^b)R_{t+1}^2 \tag{46}$$

So whether the disparity increases or decreases depends upon whether π_t^a falls less or more than π_t^b in response to a decrease in $\Delta\bar{\eta}_t$.

Lemma 4.2. *As long as group b is a minority, and the share of the Complex task in output is not very high, $\beta < \bar{\beta}_t(p, \Delta\bar{\eta}_t)$, the disparity in the next period increases with an increase in the affirmative action today.*

This brings us to the following proposition:

Proposition 6. *If conditions mentioned in Lemma 4.2 are satisfied, then patronization in period $\tau > t$, measured by the probability with which "fail" b workers are assigned to the Complex task, needs to increase even if the disparity limit set by the govt. does not change. That is, $\zeta_\tau(\pi_\tau^a, \pi_\tau^b, \Delta\bar{\eta}_t) > \zeta(\pi_t^a, \pi_t^b, \Delta\bar{\eta}_t)$.*

Thus, beginning from a Liberal hiring regime, affirmative action of the nature being discussed here, causes incentives to worsen for both groups which spills over to worsen the cost functions of the future generations. This last effect is greater for the group whose representation rate is of concern. Consequently, future generations, would observe greater disparity in representation under Laissez-faire and increase the possibility of a civil unrest. To combat it, the govt. thus has to continue to engage in affirmative action. It is observed that the probability with which group b needs to be "patronized" increases in the later periods even without any increase in the severity of the affirmative action policy. This is the sense in which the economy seems to be locked in a *patronization trap*, brought about entirely because of over-zealous policy goals.

Finally, consider the dynamic problem of the govt. Before delving into the solution of the problem, it is important to understand how this dynamic problem is different from the one considered in the previous section. When the implementation of AA under Laissez-faire caused employers to hire more "unclear" b workers, thereby changing the composition of the labor force allotted to the two tasks but keeping the factor ratio unchanged, the legacy of the current affirmative action was not carried over into the future in any form. So the dynamic problem of the govt. could be formulated as a sequence of disconnected discrete problems and the solution involved internalizing the anticipated equilibrium and choosing the target that exactly equated the loss in output on the margin to a reduction in the expected loss from the civil unrest. The scenario being dealt in this section is very different. Here the implementation of AA causes returns for the two groups to diverge. Incentives fall for both groups, but the fall might be greater for group b if it is a minority and the share of the Complex task in the output is not very high. So the dynamic problem of the govt. in this case will be a bit more involved.

Writing out the problem of the govt. in detail:

$$\max_{\{\Delta\bar{\eta}_t\}} \delta^t [Q_t - \sum_j \lambda_j C_t^j - L\gamma(\Delta\bar{\eta}_t)]$$

where

$$C_q = \pi_t \tag{47}$$
$$S = (1-\pi_t^a)(1-p) + \lambda_b \Delta\bar{\eta}_t \tag{48}$$
$$C_t^j = (\pi_{t-1}^j A_1 + (1-\pi_{t-1}^j)A_0)\int_0^{R_t^j} 2c^2 \mathrm{d}(c) \quad j = a, b \tag{49}$$
$$\pi_t^j = \gamma^j(\pi_{t-1}^a, \pi_{t-1}^b, \Delta\bar{\eta}_{t-1}, \Delta\bar{\eta}_t) \quad j = a, b \tag{50}$$

The govt. behaves like a Stackelberg leader in choosing the sequence of targets. A choice of target today affects the cost distribution and consequently the investment behavior tomorrow. Given this, the overall effect of choosing a target in period t can be decomposed as: (1) effect on period t output, (2) effect on the probability of unrest in period t, (3) effect on period $t+1$ output, (4) effect on aggregate cost of effort in period t, and finally (5) the effect on aggregate cost of effort in period $t+1$. In what follows, I discuss each of these effects.

Begin by considering effect (1). Invoking Lemma 3.1 it is straightforward to see that output and the limit set by the govt are positively related. [15] Effect (2) is also trivial given the definition of the $\gamma(\cdot)$ function. Effect (3) captures how the total cost of effort in period t is affected by AA. To evaluate it, re-write aggregate cost in the following form:

$$\sum_j \lambda_j C_t^j = \sum_j \lambda_j \int_0^{\pi_t^j} (\frac{z}{A_{t-1}^j})^{1/2} \mathrm{d}(z) \tag{51}$$

Then, $\frac{\partial C_t^j}{\partial \Delta\bar{\eta}_t} = R_t^j \frac{\partial \pi_t^j}{\partial \Delta\bar{\eta}_t} > 0$. So lower is the limit set by the govt., lower is the aggregate cost expended. This effect obtains because AA dampens incentives and few workers choose to exert effort. Next, we consider effect (4) which tries to see how the cost distribution of the next period

[15] Verify that ϕ is greater than 0 in this case: $\phi = \frac{(\pi_t^a - \pi_t^b)(1-p) - \Delta\bar{\eta}_t}{p} = \frac{\lambda_b \zeta_t (1-\pi_t^b)(1-p)}{p}$.

is affected. Using the same approach re-write the aggregate cost of the next period as:

$$\sum_j \lambda_j C_{t+1}^j = \sum_j \lambda_j \int_0^{\pi_{t+1}^j} (\frac{z}{A_t^j})^{1/2} \mathrm{d}(z) \tag{52}$$

So,

$$\frac{\partial C_{t+1}^j}{\partial \Delta \bar{\eta}_t} = -\frac{(A_1 - A_0)}{2(A_t^j)^{3/2}} \frac{\partial \pi_t^j}{\partial \Delta \bar{\eta}_t} \int_0^{\pi_{t+1}^j} z^{1/2} \mathrm{d}(z) \tag{53}$$

Thus, the lower is the limit set by the govt. in period t, the higher is the aggregate cost expended in period $t + 1$.

The final effect that needs to be analyzed discusses how the output in period $t + 1$ is affected by AA in period t. This effect is slightly more involved than effects (1)-(4) because here one needs to measure the indirect effect of lowering $\Delta\bar{\eta}_t$ on output in period $t + 1$ through the former's effect on equilibrium priors in period $t + 1$. That is,

$$\frac{\partial Q_{t+1}}{\partial \Delta \bar{\eta}_t} = \sum_j \frac{\partial Q_{t+1}}{\partial \pi_{t+1}^j} \frac{\partial \pi_{t+1}^j}{\partial \pi_t^j} \frac{\partial \pi_t^j}{\partial \Delta \bar{\eta}_t} \tag{54}$$

Of the three terms on the right of (72), $\frac{\partial \pi_t^j}{\partial \Delta \bar{\eta}_t}$ has already been derived (see (50) and (53)). From the equations describing the equilibrium in period $t + 1$, one can derive $\frac{\partial \pi_{t+1}^j}{\partial \pi_t^j}$ as:

$$\frac{\partial \pi_{t+1}^j}{\partial \pi_t^j} = (A_1 - A_0)(R_{t+1}^j)^2 \tag{55}$$

So the only term that remains to be evaluated is $\frac{\partial Q_{t+1}}{\partial \pi_{t+1}^j}$, the effect on output of a change in the rate of investment. When investment rates change, they cause the labor input to change in the two tasks in the following ways:

$$\frac{\partial C_q}{\partial \pi_{t+1}^a} = \lambda_a \tag{56}$$

$$\frac{\partial C_q}{\partial \pi_{t+1}^b} = \lambda_b \tag{57}$$

$$\frac{\partial S}{\partial \pi_{t+1}^a} = -(1-p) \tag{58}$$

$$\frac{\partial S}{\partial \pi_{t+1}^b} = 0 \tag{59}$$

Since in this hiring regime, all "pass" and "unclear" workers are being assigned to the Complex task, so a marginal change in the investment rate of either group leads to a change in the labor input in the Complex task of the amount equal to the demographic presence of each group. Also, any increase in the investment rate among a workers leads to a decrease in the "fail" a workers; so the labor force in the Simple task falls. Finally, any increase in the investment rate among the b workers would have zero net effect on the labor force in the Simple task because the probability with which "fail" b workers are hired adjusts appropriately to maintain the disparity in the rates equal to the limit set by the govt. Given these observations, the total effect on output of a change in the investment rate in group a and b is given by:

$$\frac{\partial Q_{t+1}}{\partial \pi^a_{t+1}} = (\frac{C_q}{S})^{\beta-1}\left[\beta\lambda_a - (1-\beta)\frac{C_q}{S}(1-p)\right] \gtrless 0 \quad (60)$$

$$\frac{\partial Q_{t+1}}{\partial \pi^b_{t+1}} = (\frac{C_q}{S})^{\beta-1}\beta\lambda_b > 0 \quad (61)$$

Although for group b the effect on output of increasing the rate of investment is unambiguous, it is not so the case for group a. The increase in investment rate can raise output if the contribution of the Complex task is high enough and group a is sufficiently large. But a priori, it is hard to sign the effect and I will consider both possibilities.

When output is positively related to π^a_{t+1}, then effect (5) also turns out to be positive, once we have signed all the terms in (72). Thus, the optimal choice of $\Delta\bar{\eta}_t$ should satisfy the following:

$$\frac{\partial Q_t}{\partial \Delta\bar{\eta}_t} - \sum_j \lambda_j \frac{\partial C^j_t}{\partial \Delta\bar{\eta}_t} + \delta[\frac{\partial Q_{t+1}}{\partial \Delta\bar{\eta}_t} - \sum_j \lambda_j \frac{\partial C^j_{t+1}}{\partial \Delta\bar{\eta}_t}] = L\gamma'(\Delta\bar{\eta}_t) \quad (62)$$

Collecting terms with the same sign on one side:

$$\frac{\partial Q_t}{\partial \Delta\bar{\eta}_t} + \delta[\frac{\partial Q_{t+1}}{\partial \Delta\bar{\eta}_t} - \sum_j \lambda_j \frac{\partial C^j_{t+1}}{\partial \Delta\bar{\eta}_t}] = \sum_j \lambda_j \frac{\partial C^j_t}{\partial \Delta\bar{\eta}_t} + L\gamma'(\Delta\bar{\eta}_t) \quad (63)$$

So the choice of the limit in this case is derived by equating the marginal benefits and losses. Losses involve sacrificing output today and tomorrow and a facing a higher cost tomorrow. Benefits involve lowered risk of a civil unrest and lower cost of effort expended today.

5 Discussion and Concluding Remarks

This paper set out to analyze the optimal affirmative action policy in a world of intergenerational spillovers. The general equilibrium model with imperfect information and identity-sighted regime was seen to perform below efficiency. Moreover, unequal histories were reflected in the cost of effort faced by workers in the two groups which translated to very unequal representations in the high-paying Complex task. Since it was assumed that a prolonged period of inequality might potentially cause a civil unrest to break out, a case was made for government intervention in the form of a limit on the disparity in the representation of the two groups.

The static consequences of affirmative action involved the employers devising a hiring policy that was more lenient towards members of the under-privileged group and stricter towards members of the privileged group. It was observed that beginning from hiring regimes which were biased against the discriminated group, affirmative action served to improve incentives for both groups. And it served to worsen incentives for the hiring regimes which were not very strict with the discriminated group to begin with.

The optimal policy for a period was derived by equating the temporal cost and benefit margins. Affirmative action caused output to fall but at the same time improved future cost distributions by improving incentives today. These effects were weighed against the reduced probability of a social unrest to give the optimal policy.

Effects of affirmative action were markedly different in the hiring regime that was most lenient towards the under-privileged group. It was observed that incentives would worsen unambiguously for both groups and the disparity in representation under Laissez-faire would be higher in the next period, if the target group was minority. This called for a more lenient hiring policy towards the target group to contain the disparity, which caused incentives to worsen further and so on. The result was an economy that would continue to need affirmative action for a very long time in the future. Since affirmative action were causing incentives to differ for the two groups, convergence would not be possible. The economy will settle at a much lower long run rate of investment in human capital as compared to what it would have achieved under Laissez-faire. This highlights the possibility that governments, in their zealous attempts to improve the plight of a group they consider under-privileged, might end up putting the economy on a sub-optimal path that produces entirely unwanted results.

It is important to mention the role that certain aspects of the model played in deriving these results. First, I consider a Cobb-Douglas technology which made incentives a function of the factor ratio. Consequently, when affirmative action caused only the composition of the labor input in the two tasks to change without changing the factor ratio, incentives remained unaffected. And since rates of investment in human capital was the vehicle of inter-generational externality, this ensured that the effects of affirmative action in any period remained confined to that period and did not spillover to the future generations. This scenario, was, of course, not true when affirmative action was launched at a hiring regime that was highly lenient towards the target group in the first place. Then it was observed that, affirmative action of today goes on to increase the cost of effort for future generations and might also decrease output.

Second, I assume a simple three-outcome testing technology. The analysis can of course be done with a more general testing technology with a continuum of outcomes. What I gain by sacrificing generality is a clear characterization of the possible equilibria in the labor market and this allows me to analyze the static and dynamic effects of affirmative action in greater detail. Also, my testing technology allowed generated ambiguous results for the qualified and the unqualified with the same probability, a fact that simplified the analysis to a great deal. If they had been different, there would not have been any qualitative changes in the equilibrium in the labor market, but there would have been some changes in the static effects of affirmative action in the non-patronization case. The result that incentives are not affected by the choice of the limit would not be true any longer and rates of human capital investment would depend explicitly on affirmative action policies.

Finally, I assume parental qualification is what affects a child's cost of acquiring education. However, many different factors can be identified as being sources of intergenerational externality. For example, it is possible to think of the parent's occupation as the factor affecting the next generation's ability and not parental qualification. If this is the case, then, with very slight modifications to the existing set-up, one can show that very similar conclusions obtain. As long as the economy is not already very lenient towards the target group, affirmative action causes the representation of the target group workers to improve and that of the privileged group workers to fall. This improves cost distributions for the former and worsens it for the latter in the next generation. In other words, it moves the economy towards a no-disparity long run steady state. One reason for not considering this set up is my premise that a one-shot government imposed re-assignment of workers in jobs does not really erase the entire legacy of history. Occupational status does have externality effects, but they cannot entirely explain persistent backwardness of a group.

To sum, this paper considered the question of affirmative action over time and derived conditions that define the optimal policy choice in each period. It also showed when affirmative action can cause the economy to become dependent on it. It is highly important that policy-makers consider the extent to which the target group is actually under-privileged before committing themselves to any positive discrimination policy.

References

[1] Kenneth Arrow. The theory of discrimination. Working Papers 403, Princeton University, Department of Economics, Industrial Relations Section., October 1971.

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

[3] Shubham Chaudhuri and Rajiv Sethi. Statistical discrimination with peer effects: Can integration eliminate negative stereotypes? *Review of Economic Studies*, 75(2):579–596, April 2008.

[4] Stephen Coate and Glenn C Loury. Will affirmative-action policies eliminate negative stereotypes? *American Economic Review*, 83(5):1220–40, December 1993.

[5] Elena del Rey and Mara Racionero. An efficiency argument for affirmative action in higher education. *Hacienda Pblica Espaola*, 187(4):41–48, December 2008.

[6] Partha S. Ghosh. Positive discrimination in india: A political analysis. *Ethnic Studies Report*, XV(2):145, July 1997.

[7] Glenn Loury. Conceptual problems in the enforcement of anti-discrimination laws. pages 296–316, 2000.

[8] Glenn C. Loury. *The Anatomy of Racial Inequality*. Harvard University Press, 2002.

[9] Andrea Moro and Peter Norman. Affirmative action in a competitive economy. *Journal of Public Economics*, 87(3-4):567–594, March 2003.

[10] Andrea Moro and Peter Norman. A general equilibrium model of statistical discrimination. *Journal of Economic Theory*, 114:1–30, 2004.

6 Appendix

Proof of Lemma 2.1

Proof The first order conditions of the employers' maximization are:

$$\begin{aligned}
(\psi \le 1): \quad \frac{\partial Q}{\partial C_q} &\le \frac{\partial Q}{\partial S} \\
(0 \le \alpha^a \le 1): \quad \pi^a \frac{\partial Q}{\partial C_q} &\lesseqgtr \frac{\partial Q}{\partial S} \\
(0 \le \alpha^b \le 1): \quad \pi^b \frac{\partial Q}{\partial C_q} &\lesseqgtr \frac{\partial Q}{\partial S} \\
(0 \le \zeta \le 1): \quad 0 &\le \frac{\partial Q}{\partial S}(1-\pi)(1-p)
\end{aligned}$$

Clearly, for $\zeta > 0$, we need $\frac{\partial Q}{\partial S}(1-\pi)(1-p) = 0$ which implies implies $\pi^a \frac{\partial Q}{\partial C_q} > \pi^b \frac{\partial Q}{\partial C_q} > \frac{\partial Q}{\partial S} = 0$. So α^a and α^b must equal 1.

For α^a or α^b to be strictly positive, $\pi^j \frac{\partial Q}{\partial C_q} = \frac{\partial Q}{\partial S}$, which implies $\frac{\partial Q}{\partial C_q} < \frac{\partial Q}{\partial S}$ since $\pi^j < 1$ for $j = a, b$.

If $\pi^a > \pi^b$, then $\pi^a \frac{\partial Q}{\partial C_q} > \pi^b \frac{\partial Q}{\partial C_q}$. So if $\alpha^b > 0$, $\pi^b \frac{\partial Q}{\partial C_q} \ge \frac{\partial Q}{\partial S}$ implies $\pi^a \frac{\partial Q}{\partial C_q} > \frac{\partial Q}{\partial S}$, or, α^a must equal 1.

If $\psi < 1$, then $\frac{\partial Q}{\partial C_q} = \frac{\partial Q}{\partial S}$. So clearly, $\pi^j \frac{\partial Q}{\partial C_q} < \frac{\partial Q}{\partial S}$ for any $\pi^j \in (0,1)$. Thus, $\alpha^a = \alpha^b = 0$.

$\blacksquare$.

Details about the different hiring regimes

C **Conservative**: Employers operate at C. Labor input in the Complex task equals all workers who get a "pass" signal. So effective labor input equals $\pi(1-p)$. All remaining workers are put into the Simple task. So $S = 1 - \pi(1-p)$. The slope of the isoquant is higher than π^a:

$$\pi^a < \frac{\frac{\partial Q}{\partial S}}{\frac{\partial Q}{\partial C_q}} = \frac{1-\beta}{\beta} \frac{\pi(1-p)}{1-\pi(1-p)}$$

P **Partial**: Employers operate on the DC stretch. The total labor input in the Complex task comprises of all "pass" workers and some "unclear" a workers, given by $\pi(1-p) + \lambda_a p \alpha^a$.

Effective labor input is $\pi(1-p) + \lambda_a p \pi^a \alpha^a$. All remaining workers are put in the Simple task. The slope of the isoquant must equal π^a. Thus we get (4).

(D) **Discriminatory**: Employers operate at the point D. The total labor input in the Complex task is given by all "pass" workers and all "unclear" a workers. Effective labor input is given by $\pi(1-p) + \lambda_a p \pi^a$. Remaining put in Simple task. Slope of the isoquant must be greater than π^b but less than π^a. This gives (5).

(I) **Indifferent**: Employers operate on the DL stretch. Total labor input in the Complex task comprises of all "pass" workers, all "unclear" a workers and some "unclear" b workers. Effective labor input is given by $\pi(1-p) + \lambda_a \pi^a p + \lambda_b \pi^b p \alpha^b$. The slope of the isoquant should equal π^b. This gives condition (6).

(L) **Liberal**: Employers operate at point L. All "pass" and "unclear" workers from both groups are assigned to the Complex task. The effective labor input is given by $\pi(1-p) + \pi p$. The slope of the isoquant is less than π^b. This gives (7).

Details about Returns to effort Here I provide detailed derivations about the returns to exerting effort under the different hiring regimes and show how we arrive at (10).

C **Conservative**:

$$
\begin{aligned}
R(\pi^a, \pi^b; \alpha^a = 0, \alpha^b = 0) &= (1-p)\left(\frac{\partial Q}{\partial C_q} - \frac{\partial Q}{\partial S}\right)\Big|_C \\
&= (1-p)\frac{\partial Q}{\partial C_q}\left(1 - \frac{\frac{\partial Q}{\partial S}}{\frac{\partial Q}{\partial C_q}}\right)\Big|_C \\
&= (1-p)\beta (C_q/S)^{\beta-1}\left(1 - \frac{\mathrm{d}C_q}{\mathrm{d}S}\right)\Big|_C \\
&= (1-p)\beta \left\{\frac{\beta}{1-\beta}\rho_0\right\}^{\beta-1}(1-\rho_0) \\
&= (1-p)\beta^\beta (1-\beta)^{1-\beta}\rho_0^{\beta-1}(1-\rho_0)
\end{aligned}
$$

P **Partial**:

$$\begin{aligned} R(\pi^a, \pi^b; \alpha^a > 0; \alpha^b = 0) &= (1-p)(1-\pi^a)\frac{\partial Q}{\partial C_q} \\ &= (1-p)(1-\pi^a)\beta(C_q/S)^{\beta-1} \\ &= (1-p)(1-\pi^a)\beta\left\{\frac{\beta}{1-\beta}\pi^a\right\}^{\beta-1} \\ &= (1-p)\beta^\beta(1-\beta)^{1-\beta}(\pi^a)^{\beta-1}(1-\pi^a) \end{aligned}$$

D **Discriminatory**:

$$\begin{aligned} R(\pi^a, \pi^b; \alpha^a = 1, \alpha^b = 0) &= (1-p)(\frac{\partial Q}{\partial C_q} - \frac{\partial Q}{\partial S})\big|_D \\ &= (1-p)\frac{\partial Q}{\partial C_q}(1 - \frac{\frac{\partial Q}{\partial S}}{\frac{\partial Q}{\partial C_q}})\big|_D \\ &= (1-p)\beta(C_q/S)^{\beta-1}(1 - \frac{\mathrm{d}C_q}{\mathrm{d}S})\big|_D \\ &= (1-p)\beta\left\{\frac{\beta}{1-\beta}\rho_{1,0}\right\}^{\beta-1}(1-\rho_{1,0}) \\ &= (1-p)\beta^\beta(1-\beta)^{1-\beta}\rho_{1,0}^{\beta-1}(1-\rho_{1,0}) \end{aligned}$$

I **Indifferent**:

$$\begin{aligned} R(\pi^a, \pi^b; \alpha^a = 1; \alpha^b > 0) &= (1-p)(1-\pi^b)\frac{\partial Q}{\partial C_q} \\ &= (1-p)(1-\pi^b)\beta(C_q/S)^{\beta-1} \\ &= (1-p)(1-\pi^b)\beta\left\{\frac{\beta}{1-\beta}\pi^b\right\}^{\beta-1} \\ &= (1-p)\beta^\beta(1-\beta)^{1-\beta}(\pi^b)^{\beta-1}(1-\pi^b) \end{aligned}$$

L **Liberal**:

$$\begin{aligned}
R(\pi^a, \pi^b; \alpha^a = 1, \alpha^b = 1) &= (1-p)(\frac{\partial Q}{\partial C_q} - \frac{\partial Q}{\partial S})\big|_L \\
&= (1-p)\frac{\partial Q}{\partial C_q}(1 - \frac{\frac{\partial Q}{\partial S}}{\frac{\partial Q}{\partial C_q}})\big|_L \\
&= (1-p)\beta(C_q/S)^{\beta-1}(1 - \frac{\mathrm{d}C_q}{\mathrm{d}S})\big|_L \\
&= (1-p)\beta\left\{\frac{\beta}{1-\beta}\rho_1\right\}^{\beta-1}(1-\rho_1) \\
&= (1-p)\beta^{\beta}(1-\beta)^{1-\beta}\rho_1^{\beta-1}(1-\rho_1)
\end{aligned}$$

Details about ρ_0,$\rho_{1,0}$ and ρ_1 Here I provide details about the ρ_i , $i = 0, (1,0), 1$ introduced in the text and also prove Lemma 2.2.

ρ_0 represents the slope of the isoquant at C; $\rho_{1,0}$ represents the slope of the isoquant at D and ρ_1 represents the slope of the isoquant at L. Clearly, they are functions of π^a and π^b. Some of their properties are listed below:

- $\rho_0(0,0) = \rho_{1,0}(0,0) = \rho_1(0,0) = 0$.
- $\rho_0(1,1) = \frac{1-\beta}{\beta}\frac{1-p}{p}$; $\rho_{1,0}(1,1) = \frac{1-\beta}{\beta}\frac{\lambda_a+\lambda_b(1-p)}{\lambda_b}$; $\rho_1(1,1) \to \infty$.
- ρ_0, $\rho_{1,0}$ and ρ_1 are monotonically increasing in π^a and π^b.

$$\begin{aligned}
\frac{\partial \rho_0}{\partial \pi} &= \frac{1-\beta}{\beta}\frac{1-p}{(1-\pi(1-p))^2} > 0 \\
\frac{\partial \rho_{1,0}}{\partial \pi^a} &= \frac{1-\beta}{\beta}\left[\frac{\lambda_a}{S} + \frac{C_q}{S}\lambda_a(1-p)\right] > 0 \\
\frac{\partial \rho_{1,0}}{\partial \pi^b} &= \frac{1-\beta}{\beta}\lambda_b(1-p)[\frac{1}{S} + \frac{C_q}{S^2}] > 0 \\
\frac{\partial \rho_1}{\partial \pi} &= \frac{1-\beta}{\beta}\frac{1}{(1-\pi)^2} > 0
\end{aligned}$$

- ρ_0, $\rho_{1,0}$ and $\rho_1(\cdot)$ are convex in π^a. And ρ_0, ρ_1 are convex in π^b and $\rho_{1,0}$ is concave in π^b.

$$
\begin{aligned}
\frac{\partial^2 \rho_0}{\partial \pi^2} &= 2\frac{1-\beta}{\beta}\frac{(1-p)^2}{1-\pi(1-p)}^3 > 0 \\
\frac{\partial^2 \rho_{1,0}}{\partial (\pi^a)^2} &= 2\frac{1-\beta}{\beta}\lambda_a\frac{C_q\lambda_a(1-p)}{S^3} > 0 \\
\frac{\partial^2 \rho_{1,0}}{\partial (\pi^b)^2} &= 2\frac{1-\beta}{\beta}\frac{(\lambda_b(1-p))^2(1-p\lambda_a^2(1-\pi^a))}{S^3} < 0 \\
\frac{\partial^2 \rho_1}{\partial \pi^2} &= 2\frac{1-\beta}{\beta}\frac{1}{(1-\pi)^3(1-p)} > 0
\end{aligned}
$$

Using their expressions, one can solve for values of π^a in terms of π^b and the parameters of the model and plot these expressions for π^a in a unit square along with π^b. These functions would then demarcate regions in which a particular inequality, of those listed in (8), would hold and would give sets of values of π^a and π^b for which particular hiring regimes would obtain.

Proof of Lemma 2.3 Proof: Since ρ_0, $\rho_{1,0}$ and ρ_1 are continuous and increasing functions of π^a and π^b, so $\Psi(\cdot)$ and $\Phi(\cdot)$ are also continuous and increasing functions. And from (13), since

$$\frac{\partial R}{\partial \Phi} = -K\Phi^{\beta-2}[\beta\Phi + (1-\beta)] < 0$$

it is straight forward to observe that R is a continuous and decreasing function of π^a, π^b.

■

Proof of Proposition 2.1

To formally prove the existence of the equilibrium, I define the following mappings:

$$
\begin{aligned}
\Gamma_a(\pi^a,\pi^b) &\equiv G_a(K\Phi^{\beta-1}(\pi^a,\pi^b)(1-\Phi(\pi^a,\pi^b)) \\
\Gamma_b(\pi^a,\pi^b) &\equiv G_b(K\Phi^{\beta-1}(\pi^a,\pi^b)(1-\Phi(\pi^a,\pi^b))
\end{aligned}
$$

Suppose that the proportion of qualified workers from the two groups equals π^a and π^b respectively. Employers will evaluate the ρ_i functions for $i = 0, (1,0), 1$ and derive the hiring policies, given by $\alpha^a(\pi^a,\pi^b)$ and $\alpha^b(\pi^a,\pi^b)$, according to (8). Given the hiring policies, workers will invest if their cost is not higher than $R(\pi^a,\pi^b,\alpha^a,\alpha^b)$. Thus, given a possible pair of priors, the mappings give the fraction of workers who choose to exert effort from each group. An equilibrium of this model is just a fixed point of the mappings $(\Gamma_a, \Gamma_b) : [0,1]^2 \rightarrow [0,1]^2$. The best responses of the employers and the workers are continuous and consequently the mappings are definitely upper hemi-continuous. Since the mappings are defined on the compact and convex set $[0,1]^2$, by Kakutani's fixed point theorem there exists at least one fixed point, or equilibrium pair. That this pair will be unique is because the G_a and G_b functions are non-decreasing and $R(\cdot)$ is monotonically decreasing in the priors.

■

Proof of Proposition 2

Proof To show the existence of the steady state, it is required to show that π_t^j is an increasing function of π_{t-1}^j:

$$\frac{\partial \pi_t^j}{\partial \pi_{t-1}^j} = \frac{A_1 - A_0}{1 - A_{t-1}^j 2R_t \frac{\partial R_t}{\partial \pi_t^j}} > 0$$

And it can be shown that π_t^j is a concave function of π_{t-1}^j.

Hence, there is a unique steady state given by $\pi_*^a = \pi_*^b = \pi_*$ given by solving

$$\pi_* = (\pi_* A_1 + (1 - \pi_*)A_0)[K\rho_*^{\beta-1}(1 - \rho_*)]^2$$

where ρ_* is the slope of the isoquant at the steady state.

■

Proof of Remark 1

To show $\frac{\partial \alpha^a}{\partial \Delta \bar{\eta}} > 0$ and $\frac{\partial \alpha^b}{\partial \Delta \bar{\eta}} < 0$ for given priors:

$$\begin{aligned} \frac{\partial \alpha^a}{\partial \Delta \bar{\eta}} &= \frac{1}{p}(\frac{\lambda_b \pi^b}{\pi}) + \frac{\lambda_a \beta}{p}(\frac{\pi^a}{\pi} - 1) > 0 \\ \frac{\partial \alpha^b}{\partial \Delta \bar{\eta}} &= -\frac{\lambda_a}{p}(\frac{(1-\beta)\pi^a}{\pi} + \beta) < 0 \end{aligned}$$

■

Proof of Lemma 3.1 *Proof*: The first order condition of the employers' maximization under AA,tangency condition of the employers' maximization is:

$$\pi_t = (\frac{\mathrm{d}C_Q}{\mathrm{d}S})_t = \frac{1-\beta}{\beta}(\frac{C_Q}{S})_t^{\beta-1} \tag{64}$$

which is the same condition as faced by the employers in an identity-blind regime. Consequently, the factor ratio in the equilibrium with affirmative action would be the same as it would have been in an identity-blind regime under Laissez-faire. And the identity-blind regime can be thought of as one where $\Delta\bar{\eta} = (\pi_t^a - \pi_t^b)(1-p)$, or, $\phi = 0$. Given that $\Delta\bar{\eta}_t < (\pi^a - \pi^b)(1-p)$, or, $\phi_t > 0$, one can compare the output that obtains when $\phi = 0$ and when $\phi > 0$ and verify

$$C_q^{\beta} S^{1-\beta}|_{\phi=0} > C_q^{\beta} S^{1-\beta}|_{\phi>0}$$

to establish the required. Re-write output as $Q = (C_q/S)^{\beta} S$ on both sides of the inequality and since the factor ratio is same for the two cases, the inequality reduces to proving $S(\phi = 0) > S(\phi > 0)$:

$$\begin{aligned} S(\phi = 0) &> S(\phi > 0) && (65) \\ (1-\pi_t)(1-p) + p(1-\alpha_t) &> (1-\pi_t)(1-p) + p(1 - \lambda_a\alpha_t^a - \lambda_b\alpha_t^b) && (66) \\ \alpha_t &< \lambda_a\alpha_t^a + \lambda_b\alpha_t^b && (67) \end{aligned}$$

Replacing the expressions for α_t^a and α_t^b and simplifying, the above inequality reduces to:

$$\frac{\lambda_a\pi_t^a\phi_t}{\pi_t}(1-\beta) + \lambda_a\phi_t\beta - \lambda_a\phi_t > 0 \tag{68}$$

Further simplification on the left side yields $\pi_t^a > \pi_t$ which is always true. Thus, we have proved that as ϕ increases, output falls. And an increase in ϕ is caused by a decrease in $\Delta\eta_t$. Thus, the lower the limit, the greater is the fall in output.

■

Proof of Lemma 4.1 Proof: The equation defining equilibrium prior for group a is given by:

$$\pi_t^a = (\pi_{t-1}^a A_1 + (1-\pi_{t-1}^a)A_0)(R_t^a)^2 = A_{t-1}^a(R_t^a)^2 \tag{69}$$

The effect of $\Delta\bar{\eta}_t$ is derived by differentiating (49) implicitly.

$$\frac{\partial\pi_t^a}{\partial\Delta\bar{\eta}_t} = \frac{2A_{t-1}^a R_t^a \frac{\partial R_t}{\partial\Delta\eta_t}}{1 - 2A_{t-1}^a R_t^a \frac{\partial R_t}{\partial\pi_t^a}} \tag{70}$$

Evaluating the derivatives in (50):

$$\frac{\partial R_t}{\partial \Delta\bar{\eta}_t} = \frac{\partial R_t}{\partial \rho_t}\frac{\partial \rho_t}{\partial \Delta\bar{\eta}_t} = -\rho^{\beta-1}(1-p)(1-\beta)[\frac{\beta}{\rho}+2(\beta+\rho)] \times -\frac{\pi_t\lambda_b}{S^2} > 0 \quad (71)$$

$$\frac{\partial R_t}{\partial \pi_t^a} = \frac{\partial R_t}{\partial \rho_t}\frac{\partial \rho_t}{\partial \pi_t^a} = -\rho^{\beta-1}(1-p)(1-\beta)[\frac{\beta}{\rho}+2(\beta+\rho)] \times \frac{\lambda_a S+\pi_t(1-p)}{S^2} < 0 \quad (72)$$

Thus $\frac{\partial \pi_t^a}{\partial \Delta\eta_t}$ is positive. The effect of $\Delta\bar{\eta}_t$ on π_t^b:

$$\frac{\partial \pi_t^b}{\partial \Delta\eta_t} = \frac{2A_{t-1}^b R_t^b \frac{\partial R_t^b}{\partial \Delta\eta_t}}{1-2A_{t-1}^b R_t^b \frac{\partial R_t^b}{\partial \pi_t^b}} \quad (73)$$

The derivative on the numerator continues to be given by (51). The expression for the derivative in the denominator is now given by:

$$\frac{\partial R_t}{\partial \pi_t^b} = \frac{\partial R_t^b}{\partial \rho_t}\frac{\partial \rho_t}{\partial \pi_t^b} = (1-p)\beta(\beta-1)(\frac{C_q}{S})^{\beta-2} \times \frac{\lambda_b}{S_t} < 0 \quad (74)$$

Thus $\frac{\partial \pi_t^b}{\partial \Delta\eta_t}$ is positive.

■

<u>*Proof of Lemma 4.2*</u> *Proof*: To show $\frac{\partial \pi_t^a}{\partial \Delta\eta_t} < \frac{\partial \pi_t^b}{\partial \Delta\bar{\eta}_t}$:

$$\frac{\partial \pi_t^a}{\partial \Delta\bar{\eta}_t} - \frac{\partial \pi_t^b}{\partial \Delta\bar{\eta}_t} \gtrless 0 \Leftrightarrow \frac{2A_{t-1}^a R_t^a \frac{\partial R_t^a}{\partial \Delta\bar{\eta}_t}}{1-2A_{t-1}^a R_t^a \frac{\partial R_t^a}{\partial \pi_t^a}} - \frac{2A_{t-1}^b R_t^b \frac{\partial R_t^b}{\partial \Delta\bar{\eta}_t}}{1-2A_{t-1}^b R_t^b \frac{\partial R_t^b}{\partial \pi_t^b}} \gtrless 0 \quad (75)$$

Observe that $\frac{\partial R_t^j}{\partial \Delta\bar{\eta}_t} = \frac{\partial R_t^j}{\partial \rho_t}\frac{\partial \rho_t}{\partial \Delta\bar{\eta}_t}$ for $j=a,b$. Similarly, $\frac{\partial R_t^j}{\partial \pi_t^j} = \frac{\partial R_t^j}{\partial \rho_t}\frac{\partial \rho_t}{\partial \pi_t^j}$ for $j=a,b$. Making these substitutions allows us to cancel $\frac{\partial \rho_t}{\partial \Delta\bar{\eta}_t}$ from both terms and since it was a negative term, canceling it reverses the inequality. Finally, we define $\mathcal{A}$ and $\mathcal{B}$ for notational clarity as:

$$\mathcal{A} \equiv A_t^a R_t^a \frac{\partial R_t^a}{\partial \rho_t} \quad (76)$$

$$\mathcal{B} \equiv A_t^b R_t^b \frac{\partial R_t^b}{\partial \rho_t} \quad (77)$$

and the inequality on the right-hand side of (57) becomes:

$$\frac{\mathcal{A}}{1-2\mathcal{A}\frac{\partial \rho_t}{\partial \pi_t^a}} - \frac{\mathcal{B}}{1-2\mathcal{B}\frac{\partial \rho_t}{\partial \pi_t^b}} \lesseqgtr 0 \quad (78)$$

Simplifying, the inequality becomes:

$$\frac{\partial \rho_t}{\partial \pi_t^a} - \frac{\partial \rho_t}{\partial \pi_t^b} \lesseqgtr \frac{\mathcal{B} - \mathcal{A}}{2\mathcal{B}\mathcal{A}} \tag{79}$$

$$\frac{\pi_t(1-p) + S_t(\lambda_a - \lambda_b)}{S_t^2} \lesseqgtr \frac{\mathcal{B} - \mathcal{A}}{2\mathcal{B}\mathcal{A}} \tag{80}$$

If $\lambda_b < 1/2$, then the left hand side is always positive. If the right hand side turns out to be negative, then clearly, the $>$ inequality holds. Since $\mathcal{A}$ and $\mathcal{B}$ are negative, the right hand side amounts to $|\mathcal{A}| - |\mathcal{B}|$. Observe that $A_{t-1}^a > A_{t-1}^b$ and $R_t^a < R_t^b$ because in period t AA made a hiring of b workers a priority. Finally, $|\frac{\partial R_t^a}{\partial \rho_t}| > |\frac{\partial R_t^b}{\partial \rho_t}|$. Using the actual expressions, the right hand side can be written as:

$$|\mathcal{A}| - |\mathcal{B}| = (1-p)\rho_t^{2\beta-3}[A_{t-1}^a(1-\beta)^2(\beta - 2\rho_t(1-\beta))(\beta + 2\rho_t(\beta + \rho_t)) - A_t^b\beta^2] \tag{81}$$

As long as $\beta < \frac{2\rho_t}{1+2\rho_t}$, the right hand side of (62) will be negative. Since ρ_t is a function of $p, \beta, \Delta\bar{\eta}_t$ in equilibrium, the above inequality can be solved for an upper limit on β, denoted by $\bar{\beta}_t$ as a function of $p, \Delta\bar{\eta}_t$.

Thus

$$\frac{\partial \rho_t}{\partial \pi_t^a} - \frac{\partial \rho_t}{\partial \pi_t^b} > \frac{\mathcal{B} - \mathcal{A}}{2\mathcal{B}\mathcal{A}} \tag{82}$$

which means $\frac{\partial \pi_t^a}{\partial \Delta\eta_t} < \frac{\partial \pi_t^b}{\partial \Delta\eta_t}$.

■

Proof of Proposition 4 Proof:Lemma 4.2 established that when the given conditions hold, the following is true:

$$\frac{\partial \pi_{t+1}^a}{\partial \Delta\bar{\eta}_t} < \frac{\partial \pi_{t+1}^b}{\partial \Delta\bar{\eta}_t}$$

This means that the period $t+1$ investment rate for bs falls by more than the fall in the investment rate for as. Consequently, $(\pi_{t+1}^a - \pi_{t+1}^b)(1-p) > (\pi_t^a - \pi_t^b) > \Delta\bar{\eta}_t$. Combining this with the fact that intergenerational spillovers cause $\pi_{t+1}^b > \pi_t^b$, gives:

$$\zeta_{t+1} = \frac{\pi_{t+1}^a - \pi_{t+1}^b}{1 - \pi_{t+1}^b} - \frac{\Delta\bar{\eta}_t}{(1-\pi_{t+1}^b)(1-p)} > \zeta_t \tag{83}$$

The same logic can be applied to establish the result for any $\tau > t$.

■