NBER WORKING PAPER SERIES

VALUING IDENTITY

Roland G. Fryer, Jr.
Glenn Loury

Working Paper 16568
http://www.nber.org/papers/w16568

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2010

We are grateful to Pierre-Andre Chiappori, Edward Glaeser, Lawrence Katz, Bernard Salanie and two anonymous referees for helpful comments and suggestions. We also thank seminar participants at numerous universities. Loury gratefully acknowledges financial support from the Andrew W. Mellon Foundation and the Carnegie Corporation. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Valuing Identity
Roland G. Fryer, Jr. and Glenn Loury
NBER Working Paper No. 16568
December 2010
JEL No. H2,J08,J15

ABSTRACT

Affirmative action policies are practiced around the world. This paper explores the welfare economics of such policies. A model is proposed where heterogeneous agents, distinguished by skill level and social identity, compete for positions in a hierarchy. The problem of designing an efficient policy to raise the status in this competition of a disadvantaged identity group is considered. We show that: (i) when agent identity is fully visible and contractible (sightedness), efficient policy grants preferred access to positions, but offers no direct assistance for acquiring skills; and, (ii) when identity is not contractible (blindness), efficient policy provides universal subsidies when the fraction of the disadvantaged group at the development margin is larger then their mean (across positions) share at the assignment margin.

Roland G. Fryer, Jr.
Department of Economics
Harvard University
Littauer Center 208
Cambridge, MA 02138
and NBER
rfryer@fas.harvard.edu

Glenn Loury
Department of Economics
Brown University
64 Waterman Street
Providence RI 02912
Glenn_Loury@Brown.edu

"*This is the next and the more profound stage of the battle for civil rights. We seek not just freedom but opportunity. We seek not just legal equity but human ability, not just equality as a right and a theory but equality as a fact and equality as a result.*" President Lyndon B. Johnson, Howard University, 1965

1 Introduction

When productive or developmental opportunities must be rationed in a population, the social identities of those selected can be a matter of great importance. By "identity" we refer – varying with the application – to an agent's race, sex, age, nationality, religion, ethnicity, or caste. When choosing which students to admit, employees to hire, candidates to slate, or firms to patronize, government and business actors alike often confront an intense public demand for some intervention that will engender more diversity in the ranks of the chosen. As a consequence, regulations intended to achieve this end – policies going under the rubric of "affirmative action," or "positive discrimination," or (less neutrally) "reverse discrimination" – have been promulgated in many societies throughout the world. This paper examines the welfare economics of such diversity-promoting public regulation.

Consider a few examples. In nations with sharp sectarian divisions – Lebanon, Indonesia, Pakistan, Iraq – political stability can hinge on maintaining ethnic balance in the military ranks, or on distributing coveted political offices so that no single group has disproportionate influence. In the US, selective colleges and universities often feel obliged to alter their admissions standards to enhance the racial diversity of their student bodies. Amidst rioting and civil unrest, France is struggling with how to design policies to ensure more diversity in firms. In other parts of Europe, some political parties have mandated that female candidates be adequately represented on their electoral lists. In post-Apartheid South Africa, to ensure that wealth is distributed across a wider spectrum of society a policy of "Broad Based Black Economic Empowerment" has been enacted, setting minimum numerical standards of black representation that companies are obliged to meet. In Malaysia, in the wake of widespread ethnic rioting that erupted in 1969, a "New Economic Policy" was instituted, creating quotas and preferences for ethnic Malays in public contracting, employment, and education. In India, so-called "scheduled castes and tribes" enjoy preferred access to university seats and government jobs by constitutional mandate, though amidst fierce controversy.[1] Though these varied programs differ in many details, we will refer to all such diversity-enhancing efforts as "affirmative action.[2]"

Affirmative action policies entail the preferential valuation of social identity based on a presumption that, on the average, those being preferred cannot compete on an equal basis because of a pre-existing (if not innate) social handicap. But the stubborn realities of unequal development create some unavoidable economic problems: In the short-run, at least, enhanced access for a genuinely disadvantaged group to much sought-after productive opportunities cannot be achieved without lowering standards, distorting human capital investment decisions, or both.[3] So, given that the supply of opportunity is limited, exogenous between-group differences necessarily make diversity a costly commodity. The relevant economic problem then becomes understanding how these costs should be conceptualized, and how they can be minimized.

Our analysis of the welfare economics of affirmative action policies is motivated by two thematic questions in particular:

(1) Where in the economic life cycle should preferential treatment be most emphasized – before or after productivities have been essentially determined?

(2) How do public policies that valorize a non-productive trait – i.e., identity – affect private incentives to become more productive?

To explore these questions, we develop a simple economic model of hierarchies and preferential identity valuation. A population of agents belonging to distinct social groups invest in human capital and then compete for assignments that give them an opportunity to use their skills. One

[1] Many other examples could be given, from countries such as Phillipines, Nigeria and Sri Lanka. For a comprehensive review of the use of these policies in global perspective, see Sowell (2004). On maintaining ethnic diversity in military selection, see Klitgaard(1986). On racial preferences at US colleges and universities, see Bowen and Bok (1998). On caste and ethnic preferences in India, see Galanter (1992) and Deshpande (2006).

[2] There has been much heated debate about the *fairness* of affirmative action. Yet, these policies – and the controversy they inevitably inspire – can be found virtually everywhere. For this reason, while fairness issues are a real concern, we focus on how greater racial or ethnic diversity in a hierarchy of positions can be *efficiently* achieved. For a discussion of the social justice issues raised by affirmative action and other racially egalitarian policies, see Loury (2002), Chp. 4.

[3] See Fryer and Loury (2005a) for a detailed discussion of some usually overlooked, yet unavoidable, trade-offs associated with affirmative action policies.

group is disadvantaged, and policies to enhance opportunity for the agents in that group are considered. Designing an efficient policy of this kind is posed as an elementary mechanism design problem similar to optimal taxation. Our analysis is novel, relative to the existing literature on affirmative action, in our focus on the second-best efficiency question, and in the attention we give to the *visibility* and the *timing* dimensions of these policies.[4]

The *timing* issue has to do with finding an ideal point in the developmental process to introduce a preference. We distinguish in the model between the *ex ante* and the *ex post* stages of production. Given that a productivity gap already exists, an ex post preference offers a competitive edge to less productive agents in the disadvantaged group. By contrast, an ex ante preference promotes the competitive success of the disadvantaged by fostering their prior acquisition of skills. That is, ex ante policies operate on the *development margin*; while ex post policies operate on the *assignment margin*.

The *visibility* dimension of a policy concerns an informational constraint one often encounters with affirmative action, reflected in the distinction we draw between *sighted* and *blind* policy environments. Under sightedness, assignment standards and development subsidies can be tailored to group membership at the individual level. Sighted policies are overtly discriminatory, in that otherwise similar agents from different groups are treated differently. Blind policies, in contrast, are tacitly discriminatory. They have their impact by placing a premium on some non-identity traits that are known to be more prevalent in the preferred population. Though they are facially neutral in their treatment of groups, blind preferential policies have been intentionally chosen to have group-disparate effects.

Combining these distinctions of visibility and of timing generates a 2×2 conceptual matrix that captures the main contours of affirmative action as it is practiced in the real world: Job reservations, contract set-asides, distinct admissions standards, race-normed ability tests – all exemplify sighted-ex post preferences. Instances of sighted-ex ante preferences include minority scholarship funds, group-targeted skills development programs, and costly outreach and recruitment efforts that encourage an underrepresented group to prepare for future opportunities.

[4] Earlier papers on the economics of affirmative action policies include Welch (1976), Lundberg and Startz (1983), Coate and Loury (1993), Moro and Norman (2003), and Fryer and Loury (2005b). For a review of the evidence on the effectiveness of these policies, see Holzer and Neumark (2000). For a broad policy discussion, see Fryer and Loury (2005a).

On the other hand, automatic admission for the top 10% of a state's high school classes; waiving a mandate that college applicants submit test scores; selecting among applicants partly by lot; or introducing non-identity factors that are unrelated to performance into the evaluation process – are all examples of blind-ex post preferences.[5] And, since there must be some group disparity in the distribution of endowments (otherwise, no policy promoting group equality would be needed), a blind-ex ante preference can always be put in place by subsidizing for everyone those skill-enhancing actions from which agents in a preferred group can derive the most benefit.[6]

Our analysis sheds light on the questions raised above. Take the issue of timing: If policy-making over the economic life cycle is uncoordinated, early-stage proponents of diversity may overcompensate for a group's social disadvantage by failing to take due account of subsequent efforts. Indeed, our model makes it clear that *an explicit ex ante preference is always redundant in a sighted environment when the ideal ex post preference is in place.*

Likewise, consider the issue of incentives: Economic intuition suggests that, under an efficient policy, the marginal social cost of selecting another disadvantaged agent should be the same, whether that is done by lowering productivity requirements at the assignment stage, or by raising investments in productivity at the development stage. Under sightedness, as just noted, private incentives naturally comport with this rule when ex post policies are efficient. However, under blindness, private and social returns on ex ante investments need not coincide. In fact, we can use our model to show that *private investment incentives are socially inadequate in a blind policy environment whenever members of the targeted group are relatively more likely to be found on the ex ante development margin than on the ex post assignment margin.*[7]

[5] Chan and Eyster (2003) have shown that lotteries can be used to pursue affirmative action goals when racial identity is not contractible. Fryer and Loury with Yuret (2008) generalize this result and, using data on US college admissions, go on to estimate the efficiency losses from adopting blind rather than sighted preferential policies at the ex post stage. Fryer and Loury (2005b) study blind handicapping in the context of winner-take-all tournaments.

[6] A hybrid policy environment is also conceivable – one that is sighted/ex ante but blind/ex post. The view – popular in some circles in the US – that using race in admissions to institutions of higher education is always wrong (blindness ex post), but resources can legitimately be expended to narrow a racial performance gap in secondary schools (sightedness ex ante), illustrates this hybrid approach. For arguments consistent with this hybrid view, see the work of Thernstrom and Thernstrom (1997) and (2003). One might imagine that there must be two such hybrid scenarios. Yet, we show below that ex ante visibility is irrelevant to efficient policy design when the regulator is sighted, ex post.

[7] Under this condition, then, requiring identity preferences to be both efficient and blind obligates one to promote

Although our model is simple – a continuum of positions in a hierarchy with affirmative action policies implemented via a tax/subsidy scheme – the insights gleaned from it are quite general. The key assumptions in this analysis are convex payoffs (investment in human capital can yield a positive return only if productivity exceeds some minimum threshold), and asymmetric information (the policy maker does not observe agents' endowments or productivities). Of course, as with all models, there is a limit to how much one can extrapolate from this laboratory. For instance, our result that ex ante investment subsidies are undesirable in a sighted policy environment depends critically on how one conceptualizes costs which we discuss in the concluding section.

The paper proceeds as follows. Section 2 introduces a simple model of production with investment in skills and competition for positions in a hierarchy. Section 3 brings affirmative action policy into the model to formally represent public interventions that expand opportunity for a disadvantaged group. Section 4 concludes.

2 A Job Assignment Model

Imagine a world where "agents" are assigned to "slots" to produce "widgets." These agents form a continuum of unit measure, and belong to one of two identity groups, $i \in \{a, b\}$, who compete for a continuum of differentiated jobs indexed by the closed interval $[0, 1]$. Each is endowed with a cost of effort $c > 0$, independently drawn from a probability distribution that depends on group identity. We imagine these costs to encompass everything that hinders or helps individuals invest in skills: peer and neighborhood effects, innate ability, quality of schooling, and so on. The fraction of agents in group i is denoted $\lambda_i \in (0, 1)$, with $\lambda_a + \lambda_b = 1$.

Economic activity takes place in two stages. At the ex ante stage agents, distinguished by their identity and effort cost (i, c), choose whether or not to exert effort. An example of this type of effort could be entering a training program or cracking the books in high school, where the costs can be seen as an inverse measure of the agent's endowed capacities. At the ex post stage these same agents, now distinguished by identity and productivity (i, μ), compete for access to slots, and production takes place. Here a slot might be a contract, a job, or some other scarce and remunerative professional opportunity; and productivity can be seen as a measure of the agent's

the general development of skills in the overall population! Moreover, when the stated condition fails, private investment incentives are socially excessive, and efficiency requires that ex ante skills acquisition be universally *taxed*!

acquired ability to make use of that opportunity.

The two stages of production are linked, in that productivity in a slot, $\mu \geq 0$, is taken to be a noisy function of prior effort at the individual level. Let $e \in \{0,1\}$ be an agent's effort choice, $z \in [0,1]$ denote her assigned slot, and $\beta(z)$ denote the effectiveness of the production technology at z. We assume that $\beta'(z) > 0$, which implies that no two positions are equally sensitive to worker productivity. A worker assigned to position z in the hierarchy earns a gross payoff $\mu \cdot \beta(z)$. The technology of production is such that one agent combines in fixed factor proportions with one slot to produce one widget, the market value of which depends on the productivity of that agent and the productivity of slot z.[8] Figure 1 illustrates the sequence of actions that we envision.

[FIGURE 1 GOES ABOUT HERE]

Notation and Preliminaries

The primitives of this model are the distributions of agents' costs and productivities. Let $G_i(c)$ be the probability that a group i agent is endowed with an effort cost that is less than or equal to; $G(c) \equiv \lambda_a G_a(c) + \lambda_b G_b(c)$ is the effort cost distribution for the entire population. Let $g_i(c)$ and $g(c)$ be the respective, continuous density functions. These distributions are assumed to have a common, connected support. The inverse functions, $G_i^{-1}(x)$ and $G^{-1}(x)$, $x \in [0,1]$, give the effort cost of an agent at the x^{th} quantile of the respective populations. By the Law of Large Numbers and our continuum of agents assumption, $G_i(c)$ is also the fraction of agents in group i with effort cost less than or equal to c.

We assume that group B is *disadvantaged* relative to group A, in the following sense:

$$\textit{Assumption 1:}\ \frac{g_a(c)}{g_b(c)} \text{ is a strictly decreasing function of } c.$$

Monotonicity of this likelihood ratio implies that, for c interior to the cost support: (1) $G_a(c) > G_b(c)$; (2) $\frac{G_a(c)}{G_b(c)} > \frac{g_a(c)}{g_b(c)} > \frac{1-G_a(c)}{1-G_b(c)}$; and, (3) $\frac{G_a(c)}{G_b(c)}$ and $\frac{1-G_a(c)}{1-G_b(c)}$ are both strictly decreasing functions of c.

Given effort, $e \in \{0,1\}$, let $F_e(\mu)$ be the probability that an agent's ex post productivity is no greater than μ. These distributions are also assumed to have a common, connected support. Their continuous density functions are denoted by $f_e(\mu)$. For a mass of agents with the common effort

[8] We adopt the specification of the second-stage production hierarchy introduced in Costrell and Loury (2004).

level, e, $F_e(\mu)$ is the fraction of that mass with productivity less than or equal to μ. Exerting effort raises stochastic productivity in the following sense:

$$\textit{Assumption 2}\text{:} \quad \frac{f_1(\mu)}{f_0(\mu)} \text{ is a strictly increasing function of } \mu.$$

As before, monotonicity implies, for μ interior to the productivity support: (1) $F_1(\mu) < F_0(\mu)$; (2) $\frac{F_1(\mu)}{F_0(\mu)} < \frac{f_1(\mu)}{f_0(\mu)} < \frac{1-F_1(\mu)}{1-F_0(\mu)}$; and, (3) $\frac{F_1(\mu)}{F_0(\mu)}$ and $\frac{1-F_1(\mu)}{1-F_0(\mu)}$ are both strictly increasing functions of μ.

Consider now the distribution of productivity in a population where the fraction of agents who exerted effort is π. For π and μ, define:$F(\pi,\mu) \equiv \pi F_1(\mu) + (1-\pi)F_0(\mu)$. Let $f(\pi,\mu)$ be the density and define the inverse, $F^{-1}(\pi,x)$, $x \in [0,1]$, as the productivity level at the x^{th} quantile of this ex post distribution.

Finally, let π_i be the fraction of group i agents who exert effort. Then $F(\pi_i,\mu)$ is the ex post distribution of productivity in that group, and

$$\lambda_a F(\pi_a,\mu) + \lambda_b F(\pi_b,\mu) = F(\lambda_a\pi_a + \lambda_b\pi_b,\mu) \equiv F(\pi,\mu)$$

is the corresponding productivity distribution for the population as a whole. Given Assumption 2, it is easily verified that: $\pi_a > \pi_b$ implies $\frac{f(\pi_a,\mu)}{f(\pi_b,\mu)}$ is a strictly increasing function of μ.

Direct Mechanisms

Let $Z_i(\mu)$ be an assignment policy which allocates group i workers with productivity μ to position $z_i(\mu) \in [0,1]$ and let $P_i(\mu)$ denote the associated payment function which assigns a price to each worker of group i worker announcing productivity μ. Together, (Z_i, P_i), an assignment and payment function, constitute a direct revelation mechanism. If a worker truthfully reports their productivity, then the workers payoff under mechanism (Z_i, P_i) is:

$$W_i(\mu) = \mu \cdot \beta(Z_i(\mu)) - P_i(\mu). \tag{1}$$

We now provide two definitions that help put structure on the set of potential mechanisms.

Definition 1 (Incentive Compatibility) *A direct mechanism $(Z_i(\mu), P_i(\mu))$ is incentive compatible if and only if:*

$$\mu \cdot \beta(Z_i(\mu)) - P_i(\mu) \geq \mu \cdot \beta\left(Z_i\left(\mu'\right)\right) - P_i\left(\mu'\right) \text{ for all } \mu, \mu' \text{ and all } i. \tag{2}$$

Thus, in the standard sense, a mechanism is incentive compatible if no agent has incentive to misreport her productivity. Our second definition asserts that a "balanced" mechanism is one in which total worker's payoffs equal total social surplus (as would occur under free entry). More formally:

Definition 2 (Balance) *A direct mechanism* $(Z_i(\mu), P_i(\mu))$ *is balanced if and only if*

$$\sum \lambda_i \int_{\Re_+} P_i(\mu)\, dF(\pi_i, \mu) = 0 \tag{3}$$

and group balanced if and only if 3 holds for all i.

Efficient Assignment of Workers to Slots

If there were no policy intervention of any kind, then more productive agents would out-compete the less productive for assignment to slots. So, a disadvantaged group would be underrepresented (holding a fraction of the slots that is less than their fraction in the population), since their effort costs are higher (Assumption 1), and effort is productive (Assumption 2). We refer to this null policy regime, where no attempt is made to assist the disadvantaged, as *laissez-faire (LF)*. In this scenario, $Z_i(\mu) = Z(\mu)$ and $P_i(\mu) = P(\mu)$ for each i.

Lemma 1 *Under LF, the efficient policy* $Z(\mu) = F(\pi, \mu)$.

Proof. The proof is structured in two parts. First, we show an efficient assignment is monotonic and non-decreasing. Second, we use this property to derive that $Z(\mu) = F(\pi, \mu)$. Given μ and μ', $\mu > \mu'$, the gain to switching assignment from μ to μ' should be non-positive:

$$\left(\mu' - \mu\right)\left[\beta(Z(\mu)) - \beta\left(Z\left(\mu'\right)\right)\right] \leq 0.$$

This implies $Z(\mu) \geq Z\left(\mu'\right)$ – the monotonicity of $Z(\cdot)$. Using this, we know

$$z \equiv \int_{\{\mu | Z(\mu) \leq z\}} dF(\pi, \mu) = F\left(\pi, Z^{-1}(z)\right).$$

Since $z = F\left(\pi, F^{-1}(\pi, z)\right)$, we have $F^{-1}(\pi, z) = Z^{-1}(z)$, which concludes the proof. ■

Lemma 1 shows that an efficient assignment of workers to slots is top-down; assigning workers to positions by starting from the top and allocating the more productive as yet assigned person to the highest ranked as yet unfilled positions (i.e. positive sorting).

Our first key result illustrates that the efficient policy is implementable by a unique direct, incentive compatible, and balanced mechanism.

Theorem 1 *There is a unique, incentive compatible, and direct mechanism* (Z^*, P^*) *which implements the efficient assignment:*

$$Z^*(\mu) = F(\pi, \mu) \qquad \textit{[top-down assignment]}$$

$$P^*(\mu) = \int_0^1 \left\{ \int_y^{F(\pi,\mu)} \beta'(x)\, \mu(\pi, x)\, dx \right\} dy \qquad \textit{[repositioning subsidy]}$$

Proof. By the definition of incentive compatibility, we know that a mechanism, (Z, P) is incentive compatible if and only if

$$\mu \left[\beta(Z(\mu)) - \beta\left(Z\left(\mu'\right)\right)\right] \geq P(\mu) - P\left(\mu'\right) \geq \mu' \left[\beta(Z(\mu)) - \beta\left(Z\left(\mu'\right)\right)\right] \text{ for all } \mu, \mu'. \tag{4}$$

This implies the monotonicity of $Z(\mu)$. Thus, $Z^*(\mu) = F(\pi, \mu)$ using the same logic as the proof of Lemma 1.

Let $\Delta\mu = \mu - \mu'$. Then, inequalities 4 can be written as:

$$\mu \frac{\left[\beta(Z(\mu)) - \beta\left(Z\left(\mu'\right)\right)\right]}{\Delta\mu} \geq \frac{P(\mu) - P\left(\mu'\right)}{\Delta\mu} \geq (\mu - \Delta\mu) \frac{\left[\beta(Z(\mu)) - \beta\left(Z\left(\mu'\right)\right)\right]}{\Delta\mu} \text{ for all } \mu, \mu'.$$

Taking the limit as $\Delta\mu \to$ zero, and using L'Hôspital's rule, we have:

$$\frac{dP(\mu)}{d\mu} = \mu \cdot \frac{d\beta(Z(\mu))}{d\mu}. \tag{5}$$

Using the definition of a balanced mechanism, this ODE has the following boundary condition: $\int_0^\infty P(\mu)\, dF(\pi, \mu) = 0$. (Note: $\sum_{i \in \{a,b\}} \lambda_i \cdot dF(\pi, \mu) = dF(\pi, \mu)$). With a change of variables, one can verify that $P(\mu) = \int_0^1 \int_y^{Z(\mu)} \beta'(x)\, \mu(\pi, x)\, dx dy$ solves (5). Recall: $Z(\mu)$ is uniquely determined by $F(\pi, \mu)$, which concludes the proof. ■

An important corollary follows immediately.

Corollary 1 *The unique, incentive compatible, and balanced direct mechanism,* (Z^*, P^*), *implies*

$$\frac{dW^*(\mu)}{d\mu} = \beta(F(\pi, \mu))$$

Proof. By definition of workers payoffs in (1) above, we have:

$$\frac{dW(\mu)}{d\mu} = \beta(Z(\mu)) + \mu \frac{d\beta(Z_i(\mu))}{d\mu} - \frac{dP_i(\mu)}{d\mu}.$$

Using the ODE of $P(\mu)$ and the fact that $Z(\mu) = F(\pi, \mu)$ from the proof of Theorem 1, the formula for $\frac{dW^*(\mu)}{d\mu}$ can be verified which concludes the proof. ■

Theorem 1 and Corollary 1 contain two important results. First, Theorem 1 demonstrates that the efficient allocation in Lemma 1 is implementable by a unique, direct, incentive compatible, and balanced mechanism and explicitly characterizes the mechanism. Corollary 1 is a direct result of Theorem 1 (although it requires a bit of proof) and will be useful in the proofs of results to come..

Given π and the efficient assignment $Z^*(\mu) = F(\pi, \mu)$, aggregate expost surplus is:

$$Q(\pi) = \int_0^1 \beta(z)\,\mu(\pi, z)\,dz$$

while the aggregate ex ante costs are at least:

$$C(\pi) = \int_0^\pi G^{-1}(z)\,dz.$$

Thus, social efficiency requires:

$$\pi^E = \arg\max_{\pi\in[0,1]} \{Q(\pi) - C(\pi)\}.$$

Our next result shows that the unique, incentive compatible, and balanced mechanism in a Laissez-Faire society is socially efficient.

Theorem 2 *The unique incentive compatible and balanced direct mechanism,* (Z^*, P^*), *maximizes social surplus.*

Proof. Workers invest whenever the expected benefit of investment is weakly greater than that investment. Thus, the level of ex ante investment, π^*, under the mechanism (Z^*, P^*) must satisfy:

$$\int_0^\infty W^*(\pi, \mu)\,[f_1(\mu) - f_0(\mu)]\,d\mu = G^{-1}(\pi). \tag{6}$$

Using the definition of social surplus and the fact that $Z^{*-1}(z) = F^{-1}(\pi, z) = \mu(\pi, z)$, the unique investment level that maximizes social surplus satisfies the following FOC:

$$\int_0^1 \beta(z)\,\frac{\partial \mu(\pi, z)}{\partial \pi}\,dz = G^{-1}(\pi) \tag{7}$$

Let $\Delta F = F_0(\mu) - F_1(\mu)$. Integrating by parts and using Corollary 1, one can verify:

$$\begin{aligned} \int_0^\infty W^*(\pi,\mu)\left[f_1(\mu) - f_0(\mu)\right] d\mu &= \int_0^\infty \beta\left(F(\pi,\mu)\right)\Delta F(\mu)\, d\mu \\ &= \int_0^\infty \beta\left(F(\pi,\mu)\right)\frac{\Delta F(\mu)}{f(\pi,\mu)} f(\pi,\mu)\, d\mu \end{aligned}$$

Implementing a change of variables and denoting $F(\pi,\mu) \equiv z$, we have $\frac{\partial\mu(\pi,z)}{\partial\pi} = \frac{\Delta F(\mu)}{f(\pi,\mu)}$. Putting this together:

$$\int_0^\infty \beta\left(F(\pi,\mu)\right)\frac{\Delta F(\mu)}{f(\pi,\mu)} f(\pi,\mu)\, d\mu = \int_0^1 \beta(z)\frac{\partial\mu(\pi,z)}{\partial\pi} dz.$$

Thus, equations (6) and (7) are identical, which concludes the proof. ■

The economic logic here, though straightforward, is worth highlighting explicitly: Equilibrium in the ex post slots market generates consumer's surplus for higher-productivity buyers. The increment to these rents that is anticipated from exerting effort provides the incentive for agents to incur costs at the ex ante stage. In competitive equilibrium, this private incentive equals the social return from greater effort, ensuring efficiency. These results are no great revelation, but deriving them explicitly allows us to introduce concepts that will prove useful in the analysis to follow.

3 The Simple Economics of Affirmative Action Policies

We have thus far concentrated on Laissez-Faire policies. In the efficient top-down assignment $Z^*(\mu) = F(\pi,\mu)$, the fraction of group i workers assigned to slots at or below $z \in [0,1]$ is given by $s_i^*(z) = F(\pi_i, \mu(\pi,z))$. With the disparate cost distributions described in Assumption 1, group B will be uniformly assigned to lower positions in the hierarchy:

$$s_b^*(z) - s_a^*(z) = (\pi_a - \pi_b)\cdot\Delta F\left(\mu(\pi,z)\right) > 0 \text{ for all } z \in [0,1].$$

That is, given any z, there is a greater fraction of group b workers assigned to slots at or below z. This unequal outcome is what affirmative action policies are designed to "correct." In this context, to expand "opportunity" for a disadvantaged group means to ensure that more of its members are assigned to slots further up the hierarchy. Since the laissez-faire allocation is efficient, any departure from it in the interest of identity diversity must reduce social surplus.

What group-specific assignments (beyond pure top-down) can be implemented when group identity is contractible? We begin by assuming each position is valued in a society in a specific way. Let the "status" of position z be denoted by $\gamma(z) : [0,1] \to \Re$. We assume that $\gamma(z)$ is non-decreasing and the aggregate status is bounded above and below: $\int_0^1 |\gamma(z)|\, dz < \infty$.

Definition 3 (Inequality Index) *Inequality Index, I, represents the overall status gap between two groups:*

$$I = \int_{\Re_+} \gamma(Z_a(\mu))\, dF(\pi_a, \mu) - \int_{\Re_+} \gamma(Z_b(\mu))\, dF(\pi_b, \mu)$$

The Inequality index for the efficient assignment rule, $Z^*(\mu)$, in a LF society is represented by a function of π^* :

$$\begin{aligned} I^{LF} &= \int_{\Re_+} \gamma(F(\pi^*, \mu))\, [f(\pi_a^*, \mu) - f(\pi_b^*, \mu)]\, d\mu \\ &= \left[G_a\left(G^{-1}(\pi^*)\right) - G_b\left(G^{-1}(\pi^*)\right)\right] \int_{\Re_+} \gamma(F(\pi^*, \mu))\, [f_1(\mu) - f_0(\mu)]\, d\mu. \end{aligned}$$

This leads to our formal definition of affirmative action

Definition 4 (Affirmative Action) *For any target inequality $\widehat{I} \in [0, I^{LF})$, the principal finds the level of ex ante subsidy and the mechanism (Z, P) that maximizes social surplus.*

Social scientists used to thinking about affirmative action in terms of "quotas" may doubt that this tax/subsidy formulation corresponds closely enough to real world practice. Yet, if the goal is to understand economic issues that bear on incentives and efficiency, we would argue that the fit is really quite good. Conventional duality considerations imply that any system of quantity constraints (quotas) can be associated with an implicit tax/subsidy scheme (derived from the shadow prices on those constraints) such that, if agents trade freely at tax-inclusive prices then, in the ensuing market equilibrium, an equivalent outcome obtains (up to a lump-sum redistribution of rents). For this reason, we believe that the core economic implications of policies like the ones we described in the Introduction – blind and sighted, ex ante and ex post – are adequately captured in this abstract model of regulation.

Affirmative Action in Group-Sighted Environments

Let us now consider how such policies could affect resource allocation in a sighted environment. A crucial implicit assumption that we are making in this analysis is that the regulator has no interest in redistribution as such (either between himself and the agents, or among the agents). Suppose that the group identity i is visible (or verifiable) and contractible. The principal sets the policy goal to the make the inequality index no greater than what obtains under LF. Given $s_i(z)$, the inequality index in a group-sighted environment is represented by:

$$I^{CS} = \int_0^1 \gamma(z) \left[s_b^{'}(z) - s_b^{'}(z) \right] dz$$

Recall, in our analysis of the LF society, we relied heavily on the fact that the efficient assignment rule was monotonic. Once we introduce groups and a desire for diversity, the assignment rule that satisfies the representation constraints will not be monotonic. To fill the void, we define a related notion.

Definition 5 *A group-specific assignment rule $Z_i(\mu)$ is* ***conditionally monotonic*** *if, within each group, more productive workers are assigned to higher ranked positions.*

If $Z_i(\mu)$ is conditionally monotonic, the productivity of a group i worker in position z can be expressed by $\mu|_{i,z} \equiv \mu(\pi_i, s_i(z))$. Thus, the ex post surplus is expressed by:

$$Q^{CS} = \sum_{i \in \{a,b\}} \lambda_i \int_0^1 \beta(z) \mu(\pi_i, s_i(z)) s_i^{'}(z) dz. \tag{8}$$

Notice: under the conditional monotonicity of $Z_i(\mu)$, the set of feasible assignments is isomorphic to the set of of group specific distributions of positions: $Z_i(\mu) = s_i^{-1}(F(\pi_i, \mu))$ and $s_i(z) = F\left(\pi_i, Z_i^{-1}(z)\right)$. Since all positions have to be filled, we must have $\lambda_a s_a(z) + \lambda_b s_b(z) = z$ for all $z \in [0,1]$.

Given π_a and π_b, assuming the conditional monotonicity of $Z_i(\mu)$, the ex post social surplus maximization problem can be written as:

$$\max_{s_a(z), s_b(z)} \sum_{i \in \{a,b\}} \lambda_i \int_0^1 \beta(z) \mu(\pi_i, s_i(z)) s_i^{'}(z) dz \tag{9}$$

subject to:

$$\begin{aligned}
(1)\ s_i^{'}(z) &\geq 0 \\
(2)\ \widehat{I} &\in [0, I^{LF}) \\
(3)\ z &= \sum_{i \in \{a,b\}} \lambda_i s_i(z)
\end{aligned}$$

This leads to our next result which fully characterizes affirmative action when identity is visible and contractible.

Lemma 2 *Let* $s_a^*(z), s_b^*(z), \rho^*$ *satisfy:*

$$(1) \qquad \lambda_a s_a^*(z) + \lambda_b s_b^*(z) = z$$

$$(2) \qquad [\mu(\pi_a, s_a^*(z)) - \mu(\pi_b, s_b^*(z))] = \frac{\rho^* \gamma'(z)}{\beta'(z)\lambda_a\lambda_b} = K(z)$$

$$(3) \qquad \widehat{I} = \int_0^1 \frac{\gamma'(z)}{\lambda_b}[z - s_a^*(z)]\,dz$$

$$(4) \qquad s_i^*(z) \text{ non-decreasing for all } z,\ \rho^* > 0.$$

If $K'(z) \in \left[-\left(\frac{1}{\lambda_b f_b}\right), \left(\frac{1}{\lambda_a f_a}\right)\right]$ *for all* z, *where* $f_i = f(\pi_i, \mu(\pi_i, s_i^*(z)))$, *then* $\{s_a^*(z), s_b^*(z)\}$ *is a solution to the group sighted affirmative action problem:* $\max\left\{Q^{CS}|I \leq \widehat{I}\right\}$.

Proof. Using integration by parts, Q^{CS} can be expressed by:

$$Q^{CS} = \beta(0)\cdot\overline{\mu} + \sum_{i\in\{a,b\}} \lambda_i \int_0^1 \beta'(z) \int_{s_i(z)}^1 \mu(\pi_i, x)\,dxdz,$$

where $\overline{\mu} = \sum_{i\{a,b\}} \lambda_i \int_0^1 \mu(\pi_i, x)\,dx$. Given $s_b(z) = \frac{z - \lambda_a s_a(z)}{\lambda_b}$, we can formulate the Lagrangian as follows (ignoring for now the constraint $s_i^*(z) > 0$ for all z) :

$\mathcal{L}(s_a(z), \rho) = \beta(0)\cdot\overline{\mu} + \int_0^1 \beta'(z)\left[\lambda_a \int_{s_a(z)}^1 \mu(\pi_a, x)\,dx + \int_{\frac{z-\lambda_a s_a(z)}{\lambda_b}}^1 \mu(\pi_b, x)\,dx\right]dz$
$+\rho\left(\widehat{I} - \int_0^1 \frac{\gamma'(z)}{\lambda_b}[z - s_a(z)]\,dz\right)$

From FOC with respect to $s_a(z)$, we have $\left[\mu(\pi_a, s_a(z)) - \mu\left(\pi_b, \frac{z-\lambda_a s_a(z)}{\lambda_b}\right)\right] = \frac{\rho\gamma'(z)}{\beta'(z)\lambda_a\lambda_b}$. It is straightforward to verify the second order condition holds. Thus, the solution $\{s_a^*(z), s_b^*(z)\}$ is defined by the following two equations:

$$\lambda_a s_a(z) + \lambda_b s_b(z) = z$$

and

$$\mu(\pi_a, s_a(z)) - \mu(\pi_b, s_b(z)) = K(z),$$

where $K(z) = \frac{\rho^*\gamma'(z)}{\beta'(z)\lambda_a\lambda_b}$ and ρ^* is the shadow price on the constraint $\widehat{I}$. Recall, we left out the constraint that $s_i^*(z)$ be non-decreasing for all z. By the implicit function theorem, this additional constraint can be written as:

$$\begin{bmatrix} s_a'(z) \\ s_b'(z) \end{bmatrix} = \begin{bmatrix} \frac{\partial \mu}{\partial s_a} & -\frac{\partial \mu}{\partial s_b} \\ \lambda_a & \lambda_b \end{bmatrix}^{-1} \begin{bmatrix} K'(z) \\ 1 \end{bmatrix}$$

Thus, the additional constraint is satisfied if and only if the RHS of the above matrix is weakly greater than zero.

$$RHS = \begin{bmatrix} f_a^{-1} & -f_b^{-1} \\ \lambda_a & \lambda_b \end{bmatrix}^{-1} \begin{bmatrix} K'(z) \\ 1 \end{bmatrix} = \left(\frac{f_a f_b}{\sum\limits_{i \in \{a,b\}} \lambda_i f_i} \right) \begin{bmatrix} \lambda_b & f_b^{-1} \\ -\lambda_a & f_a^{-1} \end{bmatrix} \begin{bmatrix} K'(z) \\ 1 \end{bmatrix} \geq 0.$$

This implies $s_i^*(z) \geq 0$ for all i if and only if $K'(z) \in \left[-\left(\frac{1}{\lambda_b f_b} \right), \left(\frac{1}{\lambda_a f_a} \right) \right]$, which concludes the proof. ■

Similar to Lemma 1, which described efficient policy in an LF society, Lemma 2 characterizes affirmative action policy when identity is visible and fully contractible.

Our next result, like Theorem 1, shows that there is a unique, incentive compatible, balanced, direct mechanism that implements the constrained efficient outcome when identity is visible and contractible.

Theorem 3 *There is a unique incentive compatible and group balanced direct assignment mechanism (Z_i^*, P_i^*) for all $i \in \{a, b\}$ that maximizes ex post surplus, where*

$$Z_i^*(\mu) = s_i^{*-1}(F(\pi_i, \mu))$$

$$P_i^*(\mu) = \int_0^1 \left\{ \int_y^{F(\pi_i, \mu)} \beta'\left(s_i^{*-1}(x)\right) \mu(\pi_i, x)\, dx \right\} dy.$$

Further, the solution involves no ex ante subsidy.

Proof. We begin with the first claim. The incentive compatibility constraint implies the conditional monotonicity of $Z_i^*(\mu)$, similar to the proof of Theorem 1. Therefore, solving the maximization problem in equation 9, we have the unique assignment rule $\{s_a^*(z), s_b^*(z)\}$. Conditional monotonicity implies $s_i^*(z) = F(\pi_i, \mu)$, $Z_i^*(\mu) = s_i^{*-1}(F(\pi_i, \mu))$. The group specific payment schedule follows as given above. This completes the first claim. We now show that this mechanism involves no ex ante subsidy, the proof of which is analogous to that of Theorem 2.

Ex post surplus, defined in equation 8 can be written as

$$Q^{CS} = \sum_{i \in \{a,b\}} \lambda_i \int_0^1 \beta\left(s_i^{-1}(x)\right) \mu(\pi_i, x)\, dx$$

Social surplus, S, is then:

$$S = Q^{CS} - \sum_{i \in \{a,b\}} \lambda_i \int_0^{\pi_i} G^{-1}(z)\, dz,$$

and is maximized when $\frac{dQ^{CS}}{d\pi_i} = \lambda_i G^{-1}(\pi_i)$ for all i. Using the envelope theorem, one can derive the π_i^E that maximizes social surplus:

$$\int_0^1 \beta\left(s_i^{*-1}(x)\right) \frac{\partial \mu(\pi_i, x)}{\partial \pi_i} dx = G^{-1}(\pi_i). \tag{10}$$

Now, let us check the incentive of workers under the given assignment mechanism. When there is no ex ante subsidy, the level of investment, π_i^*, is determined by:

$$\int_0^\infty W_i^*(\pi_i, \mu)\left[f_1(\mu) - f_0(\mu)\right] d\mu = G^{-1}(\pi_i), \tag{11}$$

where $W_i^* = \beta(Z_i^*(\mu)) - P_i^*(\mu)$ with $Z_i^*(\mu) = s_i^{*-1}(F(\pi_i, \mu))$. Recall, from Corollary 1, $\frac{dW_i^*}{d\mu} = \beta(Z_i^*(\mu))$. Using the same logic as the proof of Theorem 2, it is straightforward to show that equations (10) and (11) are identical for each i, which completes the proof. ■

Theorem 2 says that no *explicit* skills subsidy should be used to efficiently promote the access of a disadvantaged group to scarce positions when identity is fully contractible. Rather, taking the human capital investments in the groups as given, the regulator should simply underwrite the competitive position of disadvantaged agents at the ex post stage, to the extent necessary to meet his representation target. The anticipation of this preferential treatment provides an *implicit* subsidy to skills acquisition by the disadvantaged. And, since an ex post preference for B's raises the price of slots, it implicitly taxes the human capital acquisition of A's by lowering their anticipated return from exerting effort. In the equilibrium, this overt policy of discrimination in favor of disadvantaged agents at the ex post stage calls forth constrained-efficient levels of effort from the agents in both groups at the ex ante stage.[9]

[9] Thus, by raising the return on effort for B's and lowering it for A's, an efficient policy of sighted affirmative action must narrow the inter-group skills gap relative to *laissez-faire*. (That is, $\pi_b^s > \pi_b^*$ and $\pi_a^s < \pi_a^*$.) The phenomenon of "patronization" (see Coate and Loury, 1993), where preferential treatment undercuts incentives for the disadvantaged to acquire skills, does not occur if transfer-inclusive slots prices equate the private and social returns from effort for all agents.

Our finding here can be seen as a special case of a celebrated proposition in the theory of public finance due to Diamond and Mirrlees (1971): Absent any desire to redistribute firms' profits, a system of optimal indirect taxes avoids distortions that lead to inefficient production.[10] The ex ante versus ex post distinction in our set-up is equivalent to the Diamond-Mirrlees distinction between intermediate factors versus final consumption goods. Therefore, absent any concern on our regulator's part to redistribute rents as such, and given that his ex post intervention is efficient, no additional distortions at the ex ante stage are desirable. We stress that this would continue to be the case with a less than perfectly inelastic supply curve for slots, or with a more complex model of production than the one specified here – a model with skill complementarities across the agents, for example – even though the additive separability in (8) would not be preserved in the more complex setting.

Affirmative Action in Group-Blind Environments

We now, by contrast, consider the impact of affirmative action policy in a blind environment, where regulatory transfers cannot vary with identity. The game plan is as follows. First, we setup the general problem of redistribution in an environment where identity is not contractible. Mathematically, this is an linear programming problem in an infinite dimensional space with an infinite number of constraints (Anderson and Nash 1987). A general solution to this problem is beyond the scope of this paper, although we conclude this section with a full characterization of a simplified model where there are two jobs. Second, and most important, we provide a characterization of whether affirmative action should be focused on the development of skills or assignment of workers to slots.

To set up the problem of implementing affirmative action in group blind environments, we need a bit more notation. Let $h(z|x)$ denote the density of conditional random assignment, where x denotes productivity rank and z denotes slot rank; $h(z|x)\,dz = \Pr[z \in (z, z+dz)\,|x]$.

Definition 6 *A policy $h(z|x)$ is a feasible blind ex-post policy if the following three conditions hold:*

$$(1) \quad h(z|x) \geq 0 \quad \textit{[bi-stochastic kernel } h(z|x) \textit{ is weakly positive]}$$

[10] We thank Bernard Salanié of Columbia University for stressing the relevance to our work of this classical result.

(2) $\int_0^1 h(z|x)\,dz = 1$ *for all* x [*all workers are assigned to a position*]

(3) $\int_0^1 h(z|x)\,dx = 1$ *for all* z [*all positions filled by some worker*]

Let $\overline{\beta}(x)$ denote the expected productivity parameter of position to which worker of rank x is assigned and let $\overline{\gamma}(x)$ denote the expected status of the position to which a worker of rank x is assigned. In symbols: $\overline{\beta}(x) \equiv \int_0^1 h(z|x)\,\beta(z)\,dz$ and $\overline{\gamma}(x) \equiv \int_0^1 h(z|x)\,\gamma(z)\,dz$.[11] Thus, the optimal group-blind policy solves:

$$\max_{\{h(z|x)\}} \left\{ \int_0^1 \overline{\beta}(x)\,\mu(\pi, x)\,dx \right\} \tag{12}$$

subject to the following three constraints:

(1) $h(z|x)$ feasible;

(2) $\beta'(x) \geq 0$; and

(3) $\int_0^1 \overline{\gamma}(x)\,\frac{d}{dx}\left\{F(\pi_a, \mu(\pi, x)) - F(\pi_b, \mu(\pi, x))\right\}dx \leq \widehat{I}$.

As stated earlier, an analytical solution to 12 is beyond the scope of this paper. Yet, we do not need a general solution to understand some qualitative properties of the equilibrium. In particular, we are most interested in this paper on the optimal timing of affirmative action in a blind environments rather than the precise structure of such policies which are treated in Fryer, Loury, and Yuret (2008), Fryer and Loury (2005), and below for an example in the two slot version of the current model.

Let $\delta(\pi_a, \pi_b; x) = \left[\frac{f(\pi_a,\mu) - f(\pi_b,\mu)}{f(\pi,\mu)}\right]_{\mu=\mu(\pi,x)}$. Rewriting (3) yields:

$$\widehat{I} \geq \int_0^1 \overline{\gamma}(x)\,\delta(\pi_a, \pi_b; x)\,dx.$$

Now, consider the following thought experiment. If a blind ex post policy is being followed, when is it optimal to subsidize ex ante investment for everyone? This is the subject of our next result.

Theorem 4 *If a blind ex post policy is being implemented, then a subsidy ex ante is desirable if and only if*

$$\int_0^1 \Phi(x) \cdot \left(\frac{f(\pi_b, \mu)}{f(\pi, \mu)}\right)_{\mu=\mu(\pi,x)} dx < \left(\frac{g_b(c)}{g(c)}\right)_{c=G^{-1}(\pi)} \tag{13}$$

[11] A policy $h(z|x)$ can be implemented in an incentive compatible way if and only if it is stochastically monotonic ($\beta'(x) \geq 0$) and transfers that induce truth telling satisfy $\overline{P}(x) = \int_0^1 \int_y^x \beta^{-1}(z)\,\mu(\pi, z)\,dz\,dy$.

and

$$\int_0^1 \Phi(x)\,dx = 1,$$

where $\Phi(x) = \frac{\overline{\gamma}'(x)\Delta F(\mu(\pi,x))}{\int_0^1 \overline{\gamma}'(z)\Delta F(\mu(\pi,z))dz}$, *otherwise, it is desirable to tax ex ante human capital investment.*

Proof. A blind ex ante subsidy is desirable if and only if it reduces status inequality at the margin, taking into account the ex ante policy is blind. So, $G_a^{-1}(\pi_a) = G_b^{-1}(\pi_b) = G^{-1}(\pi)$ or $\pi_i = G_i\left(G_i^{-1}(\pi_i)\right)$. Hence, a universal subsidy is desirable if:

$$\frac{d}{d\pi}\left\{\int_0^1 \overline{\gamma}(x)\,\delta\left(G_a\left(G^{-1}(\pi)\right), G_b\left(G^{-1}(\pi)\right); x\right)dx\right\} < 0,$$

otherwise it is optimal to tax ex ante human capital investments. Integrating by parts and simplifying, we have

$$\begin{aligned} I(\pi) &= \int_0^1 \overline{\gamma}(x)\,\delta(\pi_a, \pi_b; x)\,dx \\ &= \int_0^1 \overline{\gamma}'(x)(\pi_a - \pi_b)\,\Delta F(\mu(\pi,x))\,dx \end{aligned}$$

where $\Delta F(\mu(\pi,x)) = F_0(\mu) - F_1(\mu) \geq 0$. Another simplification, yields

$$I(\pi) = \int_0^1 \overline{\gamma}'(x)\,\Delta F(\mu(\pi,x))\left[G_a\left(G^{-1}(\pi)\right) - G_b\left(G^{-1}(\pi)\right)\right]dx$$

which is the status inequality induced by (arbitrary) blind ex post assignment $h(z|x)$ if the ex ante investment rate is π in the aggregate. It is sufficient to show that the conditions in the Theorem imply $\frac{dI(\pi)}{d\pi} < 0$.

Using the facts that $\frac{d}{d\pi}(\Delta F(\mu(\pi,x))) = (f_0(\mu) - f_1(\mu))_{\mu=\mu(\pi,x)} \cdot \frac{d}{d\pi}\mu(\pi,x)$ and $\frac{d}{d\pi}\mu(\pi,x) = \left(\frac{\Delta F(\mu)}{f(\pi,\mu)}\right)_{\mu=\mu(\pi,x)}$, we can write:

$$\begin{aligned} \frac{dI}{d\pi} &= \left(\frac{g_a(c) - g_b(c)}{g(c)}\right)_{c=G^{-1}(\pi)} \cdot \left(\int_0^1 \overline{\gamma}'(x)\,\Delta F(\mu(\pi,x))\,dx\right) \\ &\quad + \left(\int_0^1 \overline{\gamma}'(x)\left[\pi_a - \pi_b\right]\left(\frac{f_0(\mu) - f_1(\mu)}{f(\pi,\mu)}\right)_{\mu=\mu(\pi,x)} \Delta F(\mu(\pi,x))\,dx\right) \end{aligned}$$

Hence, $\frac{dI(\pi)}{d\pi} < 0$ if and only if

$$\left(\frac{g_a(c) - g_b(c)}{g(c)}\right) < \int_0^1 \Phi(x)\left(\frac{f(\pi_a,\mu)}{f(\pi,\mu)} - \frac{f(\pi_b,\mu)}{f(\pi,\mu)}\right)_{\mu=\mu(\pi,x)},$$

where $\Phi(x) = \frac{\overline{\gamma}'(x)\Delta F(\mu(\pi,x))}{\int_0^1 \overline{\gamma}'(z)\Delta F(\mu(\pi,z))dz}$. Using this inequality and some tedious algebra completes the proof. ■

Theorem 4 provides conditions in which it makes sense to tax versus subsidize ex ante human capital investments in order to increase diversity when group identity is either not visible or contractible. Inequality 13 provides the basic economics of the problem. The right hand side is the difference between represented on the *development margin.* The left hand side of the inequality is the weighted mean presence of Bs represented on the *assignment margin.* When the mean presence of Bs on the assignment margin is less than the presence of Bs on the development margin, a universal skills subsidy is the optimal group-sighted affirmative action policy.

Group-Blind Affirmative Action in A Two-Job Example

One of the modelling assumptions that complicates our analysis in the group-sighted case is the continuum of assignment margins. If we were to restrict attention to two jobs, the analysis gets much simpler. Let θ denote the supply of slots available where $\gamma(z) = \gamma_0$ if $z \leq \theta$ and $\gamma(z) = \gamma_1$ if $z > \theta$.

Rather than a measure of status inequality, in this simple model Assumption 1 will ensure that there are a smaller fraction of *Bs* assigned to jobs relative to *As*. Now, suppose that a central authority (the *regulator)* wants to raise the fraction of group B agents who acquire slots to some target level, $\rho_b \in (\rho_b^*, \theta]$. We envision the following sequence of events in the two-job model: The regulator commits to a policy. Agents receive their endowments and make their effort choices. Productivities are realized. A slots market equilibrium is reached. Then, widgets are produced and their values earned. Our problem is to characterize a *constrained-efficient* affirmative action policy – one that is anticipated to achieve the regulator's representation target, $\rho_b > \rho_b^*$, in a surplus-maximizing manner.

We must first be more specific about what we mean by an "affirmative action policy." Assume that the regulator observes the agents' actions, but not their costs or productivities. His interventions are therefore limited to mandating transfers – to or from the agents – that depend on whether effort is exerted, and on whether a slot is held. Assume further that the regulator can impose an upper bound on the price at which slots trade in the ex post market, and that he requires random rationing among willing buyers whenever there is excess demand. Then an *affirmative action policy* is simply a list of the (action×identity)-contingent transfers, and the upper bound on the price of slots, to which the regulator is committed. This policy is blind if its transfers do not depend on identity.

Since any uniform transfer to slot holders gets capitalized into the slots price, the only way a blind regulator can promote B's access to slots at the ex post stage is to hold the price down below its market clearing level, and randomly ration slots among agents willing to buy at this artificially low price. This brings in agents just below the assignment margin, and excludes an equal number of infra-marginal agents. By Assumption 2, the former are more likely to be disadvantaged than are the latter. So, a blind ex post preference lowers productivity standards by moving from "select the very best" to "select those who are good enough." The more modest is the "good enough" threshold, the higher will be the rate at which B's get assigned to slots.

Blind policy at the ex ante stage works by imposing an across-the-board tax or subsidy on effort. This either narrows or widens, respectively, the set of agents who exert effort. How such a policy affects slot assignment rates turns on the relative numbers in the two groups of agents who are at or near the development margin. If many more B's than A's have costs just a bit too high to warrant exerting effort under *laissez faire*, then an effort subsidy will cause the equilibrium rate at which B's hold slots to rise. Conversely, if B's are comparatively scarce on the development margin, then a tax suppresses the effort of A's by more than that of B's and, in this way, causes the rate at which B's hold slots to rise in equilibrium.

A blind affirmative action policy is fully determined in this model by the two numbers – (σ, p^c) – which denote, respectively, the regulator's transfer to all agents who exert effort, and his binding upper limit on the price of slots. Let p^c be a binding $[1 - F(\pi, p^c) > \theta]$ price ceiling. With random rationing, the probability, α^c, that a willing buyer $[\mu \geq p^c]$ is assigned to a slot is:

$$\alpha^c \equiv \frac{\theta}{1 - F(\pi, p^c)} < 1. \tag{14}$$

Now, let some blind policy (σ, p^c) be given. Since agents choose effort rationally, and since under blindness agents from both groups anticipate the same returns, the equilibrium induced by this policy must have effort rates satisfying:

$$\pi = G(\sigma + \alpha^c \int_{p^c}^{\infty} \Delta F(\mu) d\mu) \tag{15}$$

and

$$\pi_i = G_i(G^{-1}(\pi)), \quad i = a, b. \tag{16}$$

Moreover, B's are holding slots at rate $\rho_b = \alpha^c \cdot [1 - F(\pi_b, p^c)]$ in this equilibrium. Therefore, using (14) and (16), the blind regulator's representation constraint can be written as:

$$0 = \rho_b \cdot [1 - F(\pi, p^c)] - \theta \cdot [1 - F(G_b(G^{-1}(\pi)), p^c)] \equiv R(\pi, p^c; \rho_b). \tag{17}$$

Under our assumptions $R(\pi, p^c; \rho_b)$ strictly increases with p^c.[12] So, with π fixed, for every $\rho_b \in (\rho_b^*, \theta]$ there is a unique price ceiling p^c which allows the blind regulator to achieve his representation target. And, the more aggressive is the target, the lower is this price ceiling.

Finally, notice that under blindness social surplus is given, once an aggregate effort rate (π) and a price ceiling (p^c) are specified: The population productivity distribution and aggregate effort costs are functions of π; and agents with$\mu \geq p^c$ have the probability $\alpha^c = \frac{\theta}{1-F(\pi,p^c)}$ of being assigned a slot, so the aggregate of widget values is also given. It follows that a blind regulator's problem can be recast as choosing π and p^c to maximize social surplus, subject to the constraint (17). Given his choice of (π, p^c), (14) and (15) can then be used to find the unique ex ante transfer, σ, which induces the chosen effort rate to arise in equilibrium. Indeed:

$$\sigma = G^{-1}(\pi) - [\frac{\theta}{1 - F(\pi, p^c)}] \cdot \int_{p^c}^{\infty} \Delta F(\mu) d\mu \tag{18}$$

We have arrived at a potentially useful characterization of ideal policy:

Lemma 3 *In a blind environment, let $(\tilde{\pi}, \tilde{p}^c)$ be the population effort rate and slots price ceiling that maximize social surplus subject to the constraint (17). Then the efficient affirmative action policy is $(\tilde{\sigma}, \tilde{p}^c)$, for $\tilde{\sigma}$ given by (18) with $\pi = \tilde{\pi}$ and $p^c = \tilde{p}^c$.*

---

[12] To prove monotonicity, notice that:

$$\frac{d}{dp^c} R(\pi, p^c; \rho_b) = \theta f(\pi_b, p^c) - \rho_b f(\pi, p^c) \lesseqgtr 0 \text{ as } \frac{f(\pi_b, p^c)}{f(\pi, p^c)} \lesseqgtr \frac{\rho_b}{\theta} = \frac{1 - F(\pi_b, p^c)}{1 - F(\pi, p^c)},$$

using (17). But, by Assumption 1 and (16), $\pi_b < \pi$ in any blind equilibrium. Assumption 2 then implies

$$\frac{f(\pi_b, p^c)}{f(\pi, p^c)} > \frac{1 - F(\pi_b, p^c)}{1 - F(\pi, p^c)},$$

which establishes the result.

A blind regulator in effect chooses an aggregate effort rate, π, and a price ceiling, p^c, to maximize surplus subject to his representation constraint (17). The efficient blind affirmative action policy $(\tilde{\sigma}, \tilde{p}^c)$ is then defined by (18).

Recall from (17) that in a blind policy environment the representation constraint is: $R(\pi, p^c; \rho_b) = 0$. Moreover, since the measure θ of slots is assigned at random to agents with $\mu \geq p^c$, social surplus under blindness is just θ times the mean productivity of these willing buyers when the population effort rate is π, net of the effort costs. That is:

$$\text{blind surplus} \equiv S(\pi, p^c) = \theta\mu^+(\pi, p^c) \; - \; \int_0^\pi G^{-1}(z)dz, \tag{19}$$

where

$$\mu^+(\pi, p^c) \equiv E[\mu \mid \mu \geq p^c; \pi] = [1 - F(\pi, p^c)]^{-1} \int_{p^c}^\infty \mu f(\pi, \mu)d\mu$$

is the conditional mean productivity among willing buyers at the slots price p^c.

Thus, by Lemma 2, the efficient blind policy is defined by the optimization problem:

$$\max_{(\pi, p^c)} \{S(\pi, p^c) \mid R(\pi, p^c) = 0\}.$$

A necessary condition for $(\tilde{\pi}, \tilde{p}^c)$ to solve this problem is that:

$$[\frac{\partial S/\partial \pi}{\partial R/\partial \pi}](\tilde{\pi}, \tilde{p}^c) = [\frac{\partial S/\partial p^c}{\partial R/\partial p^c}](\tilde{\pi}, \tilde{p}^c). \tag{20}$$

This says that the loss of surplus from raising the rate at which group B agents gain access to slots should be the same at the margin, whether this increased representation of B's is accomplished by shifting the ex ante effort rate, π, or by lowering the ex post productivity threshold, p^c.

Carrying out the tedious differentiation in (20) and rearranging terms, the first-order condition which, together with (17), defines the efficient blind policy can be written as:

$$G^{-1}(\pi) \; - \; \frac{\theta}{1 - F(\pi, p^c)} \cdot \int_{p^c}^\infty \Delta F(\mu)d\mu \; = H(\pi, p^c) \cdot K(\pi, p^c), \tag{21}$$

for functions H and K defined as follows:

$$H(\pi, p^c) \equiv \frac{\theta \Delta F(p^c)}{1 - F(\pi, p^c)} \cdot [\mu^+(\pi, p^c) - p^c] > 0,$$

and[13]

$$K(\pi, p^c) \equiv \frac{\frac{g_b}{g} - \frac{f_b}{f}}{\frac{f_b}{f} - \frac{\rho_b}{\theta}} \gtreqless 0 \;\text{ as }\; \frac{g_b}{g} \gtreqless \frac{f_b}{f}.$$

[13]The inequality here relies on the fact that, under our assumptions: $\frac{\rho_b}{\theta} = \frac{1-F(\pi_b,p^c)}{1-F(\pi,p^c)} < \frac{f_b}{f}$.

Here we are employing the notational short-hand:

$$\frac{f_b}{f} \equiv \frac{f(G_b(G^{-1}(\pi)), p^c)}{f(\pi, p^c)} \quad \text{and} \quad \frac{g_b}{g} \equiv \frac{g_b(G^{-1}(\pi))}{g(G^{-1}(\pi))}.$$

Notice that $\frac{\lambda_b f_b}{f}$ is the relative number of B's among agents on the ex post assignment margin (i.e., with $\mu = p^c$). And $\frac{\lambda_b g_b}{g}$ is the relative number of B's on the ex ante development margin (i.e., with $c = G^{-1}(\pi)$). So, comparing (21) with (18) in the light of Lemma 2 establishes the following result:

Proposition 1 *Given a representation target $\rho_b \in (\rho_b^*, \theta]$, let $(\tilde{\pi}, \tilde{p}^c)$ satisfy (21) and (17). Then, the efficient blind affirmative action policy entails the ex post ceiling price, $\tilde{p}^c$, and the ex ante effort-contingent transfer, $\tilde{\sigma}$, such that:*

$$\tilde{\sigma} = H(\tilde{\pi}, \tilde{p}^c) \cdot K(\tilde{\pi}, \tilde{p}^c).$$

The efficient policy subsidizes human capital investment for all agents ($\tilde{\sigma} > 0$) if and only if disadvantaged agents are more prevalent on the development margin than on the assignment margin.

Under laissez-faire, because the agents anticipate a market-clearing price for slots, $p^m = F^{-1}(\pi^m, 1-\theta)$, they make socially efficient human capital investment decisions. With an identity-blind preference at the ex post stage, however, agents see a lower than market-clearing price ($p^c < p^m$), and face the uncertainty of random rationing ($\alpha^c < 1$). When greater ex ante effort allows the representation target to be met with less rationing ex post, private investment incentives will be socially inadequate. This is because agents do not capture the benefits from raising their common probability of obtaining a slot (α^c) which is, in effect, a public good. When the marginal investor is more likely to belong to the disadvantaged group than is the marginal producer, a greater ex ante effort rate allows the representation target to be met in such a way that α^c rises.[14] So, this is the circumstance under which an efficient, blind affirmative action policy will provide a universal subsidy for the acquisition of skills.

[14] Using (14) and (17), and doing the differentiation, one easily shows that:

$$\frac{d\alpha^c}{d\pi}\,|_{dR=0} = \frac{\alpha^c \Delta F(p^c)}{1 - F(\pi, p^c)} \cdot K(\pi, p^c) \gtrless 0 \text{ as } \frac{g_b}{g} \gtrless \frac{f_b}{f}.$$

It is interesting to reflect on the circumstances when this critical condition – that the disadvantaged are relatively more prevalent among marginal investors than marginal producers – might be expected to hold. If slots are in especially short supply ($\theta \approx 0$), the marginal producer will fall far in the right tail of the productivity distribution where (by Assumption 2) disadvantaged agents are relatively scarce. On the other hand, when the investments needed to enhance productivity in a given pursuit are especially difficult for disadvantaged agents to make ($G_b(c) << G_a(c)$), we would expect them to be relatively scarce among marginal investors. These observations suggest that blind affirmative action should employ a general subsidy for skills acquisition only if the type of opportunities in which greater diversity is being sought are not too rarefied (i.e., difficult to qualify for), but also not too plentiful.

4 Concluding Remarks

We have shown that, when identity is fully visible and contractible, the efficient affirmative action policy avoids explicit human capital subsidies for the disadvantaged. This seemingly counter-intuitive result follows immediately, once the problem has been placed within an optimal taxation framework. To prefer a group at the assignment stage of the production process is already to implicitly subsidize their acquisition of skills. If these implicit benefits are correctly anticipated by the agents, and if the ex post intervention is itself efficient, then no further interference with investment incentives is desirable.

This is no longer the case when preferential policies must be identity-neutral. Even with an efficient ex post policy, private and social returns from ex ante investment will generally not coincide under blindness. We have shown that the second-best, identity-neutral intervention to increase opportunity for a disadvantaged group involves setting a lower productivity standard for assigning agents to slots than would occur under *laissez-faire*. Moreover, we have derived an empirically meaningful condition under which the efficient blind policy entails a general subsidy to human capital investments.

Under laissez-faire there's a common standard but not equal reputations, since Bs are on average worse than As, even given a common standard, because they invest less. Under sightedness there are two separate standards, and we compare conditional means among Bs and As above their separate standards, so the disparity has to be even greater than LF, holding investment rates constant. However, sighted AA raises incentives for Bs to invest and lowers for As, closing the gap and therefore improving Bs relative reputations. Which effect is bigger determines the ultimate impact on relative B reputation of moving from LF to sightedness: the lower standard hurts Bs relative reputation and the converging investment rates help it. Under blindness, by contrast, if investment rates didn't change from LF it would only be that a common but lower standard was being applied. We'd still be comparing means above it where Bs invest at a lower rate than as, so Bs are still relatively worse than As, but they may be "less worse" than under LF because it's a lower standard and we've made Assumption 1. On the other hand, blindness lowers investment rate in both groups (I think!) and it's not clear apriori which groups goes down by the most. If the AA goal is modest then this investment rate affect will be small. Then we could say definitively that blind AA improves Bs relative reputation relative to LF, and Bs relative reputation under LF is better than under sighted AA. Therefore, we can conclude that for a modest representation target, Bs reputation relative to As is its better under blind AA than under sighted AA. But, as noted, the relative cost of AA under blindness is much greater than under sightedness for a modest target.

References

[1] Bowen, William and Derek Bok (1998) *The Shape of the River.* Princeton: Princeton University Press.

[2] Deshpande, Ashwini (2006) "Affirmative Action in India and the United States," *2006 World Development Report.* Washington D.C.: The World Bank).

[3] Chan, Jimmy and Erik Eyster (2003) "Does Banning Affirmative Action Lower College Student Quality?" *American Economic Review* 93:2, pp. 858-72.

[4] Coate, Stephen and Glenn C. Loury (1993) "Will Affirmative Action Policies Eliminate Negative Stereotypes?" *American Economic Review,* 83:5, pp. 1220-40.

[5] Costrell, Robert and Glenn C. Loury (2004) "The Distribution of Ability and Earnings in a Hierarchical Job-Assignment Model," *Journal of Political Economy,* vol. 112, No. 6, pp 1322-

[6] Diamond, Peter A. and James A. Mirrlees (1971) "Optimal Taxation and Public Production: I – Production Efficiency," *The American Economic Review*, 61:, pp. 8-27.

[7] Fryer, Roland and Glenn C. Loury (2005a) "Affirmative Action and Its Mythology," *Journal of Economic Perspectives* 19:3.

[8] Fryer, Roland and Glenn C. Loury (2005b) "Affirmative Action in Winner-Take-All Markets," *Journal of Economic Inequality,* 3:3 pp. 263-280.

[9] Fryer, Roland and Glenn C. Loury with Tolga Yuret (in press), "Color-Blind Affirmative Action," *Journal of Law, Economics and Organization.*

[10] Galanter, Mark (1992) *Competing Equalities: Law and the Backward Classes in India.* Oxford: Oxford University Press.

[11] Holzer, Harry and David Neumark (2000) "Assessing Affirmative Action," *Journal of Economic Literature,* 38:3, pp. 483-568.

[12] Klitgaard, Robert (1986) *Elitism and Meritocracy in Developing Countries.* Baltimore: Johns Hopkins University Press.

[13] Loury, Glenn C. (2002) *The Anatomy of Racial Inequality.* Cambridge, MA: Harvard University Press.

[14] Lundberg, Shelly and Richard Startz (1983) "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review,* 73:3, pp. 340-47.

[15] Moro, Andrea and Peter Norman (2003) "Affirmative Action in a Competitive Economy," *Journal. of Public Economics,* 87:3-4, 567-94.

[16] Sowell, Thomas (2004) *Affirmative Action Around the World: An Empirical Study.* New Haven: Yale University Press.

[17] Thernstrom, Abigail and Stephen Thernstrom (1997) *America in Black and White: One Nation, Indivisible.* New York: Simon & Schuster.

[18] Thernstrom, Abigail and Stephen Thernstrom (2003) *No Excuses: Closing the Racial Gap in Learning.* New York: Simon and Schuster.

[19] Welch, Finis (1976) "Employment Quotas for Minorities," *Journal of Political Economy,* 84:4 (pt. 2), pp. S105-S139.

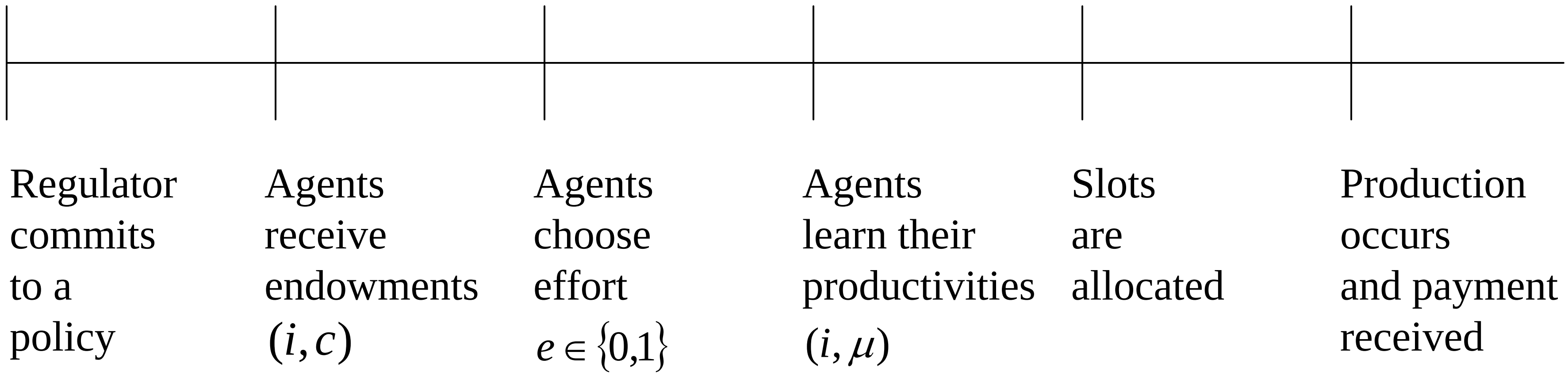

Figure 1: Sequence of Actions