Mentoring and the Dynamics of Affirmative Action

Michèle Müller-Itten * and Aniko Öry †

August 18, 2015

Abstract

We provide a dynamic overlapping generations framework to analyze the costs and benefits of affirmative action policies if mentoring complementarities are present. Senior workers reduce the young population's educational cost through mentoring, and this boost is increasing in the fraction of mentors from the same group. In such a framework, the main trade off is between using the most able workers and making the optimal use of the mentoring complementaries. We identify a large potential for market intervention in improve the long-term welfare and show that the optimal policy includes persistent regulation rather than radical short-term policies. We also contrast different channels through which the planner can affect market outcomes, and highlight the benefits of educational fellowships over hiring restrictions.

*University of California, Berkeley (email: michele.muller@berkeley.edu). Financial support from NSF (grant # SES 1227707) is gratefully acknowledged.

†Yale University (email: aniko.oery@yale.edu)

1 Introduction

For decades, affirmative action has remained a topic of heated debate and wildly opposing opinions. While some view it as an outdated form of discrimination against historically favored social classes, others see in it the only way of confronting the reality that race and gender still matter for educational achievement and employment outcomes (see e.g. Sotomayor (2014)). Part of the political discourse focuses on arguments of justice and righting historical wrongs, topics which cannot be addressed through economic research. However, the economic angle of *productivity tradeoffs* is least as relevant for policy decisions.

There is empirical evidence that race and gender affect an individual's career prospects irrespective of his or her innate ability (Milkman et al., 2014; Ellison and Swanson, 2009). Such differences in hiring rates may arise through *preference bias*, where decision makers innately favor members of a certain groups, or it may come through *statistical discrimination*, where rational utility-maximizers infer imperfectly observable information on productivity from correlated, but utility-irrelevant characteristics such as race or gender (for a comprehensive survey, see Fang and Moro (2011)). In particular, certain groups may be stuck in an equilibrium with little skill investment and poor employment prospects, while more fortunate types owe their high returns to education to favorable equilibrium beliefs.

We here focus on a third channel that comes from complementarities among productive members of the same type. Specifically, we consider mentorship relations between senior professionals and young students. It has been empirically shown that mentoring relationships are stronger and more common between members of the same demographic group (Ibarra, 1992; Dreher and Cox Jr., 1996) and that the availability of similar role models affects the academic performance of individuals (Carrell et al., 2010). In such a world, an uneven composition of the senior workforce may result in tangible productivity differences among junior workers that affect schooling and employment decisions.

This paper considers a model that incorporates the trade-off between the strong mentoring complementarities in a homogeneous work force and the optimal use of innate ability that is common to workers of all types. Our results can be applied to minorities of any type, be that based on gender, race, disability status or other demographic characteristics. Specifically, we consider an overlapping generations model where the old generation mentors the young cohort that lowers the cost of education. This mentoring boost differs across genders and is increasing in the share of mentors of one's own type. Beyond the previously mentioned evidence that mentoring relationships are commonly formed between people of similar demographics, this also captures the idea that mentors who are exposed to a diverse environment are more adept at crossing cultural barriers.

Due to the lack of suitable mentors, minority students face more hurdles in their education and are less likely to obtain a degree than their peers of equal ability. This affects their employment prospects in the unsaturated labor market during the latter part of their professional life. We assume that firms are either too small or myopic in order to internalize the long-run mentorship dynamics, and thus care only about short-term productivity without regard to the fact that today's hires affect tomorrow's employee pool.

The properties of the mentoring boost function determine the dynamics of this system, including steady state properties and per-generation productivity. We show that both a completely homogeneous and a balanced work force can be stable steady states of the economy. A homogeneous work force makes maximal use of the mentoring complementarities, while a balanced base of mentors optimally promotes the innate ability of all workers and generally increases the total labor force by boosting more workers above the industry productivity standards.

The goal of the paper is to develop conditions under which regulations can be used to increase total productivity of the economy and to compare the effectiveness of different policies. In particular, we are interested in the following questions: Under which conditions should we regulate at all? How strong should quotas or subsidies be? Is it better to have a very radical policy initially and then terminate the regulation? Who are the losers and winners of different policies?

Our model highlights that affirmative action is only useful if the following conditions are satisfied: First, the marginal benefit of more similar mentors should be significantly larger for minorities than for majorities. Furthermore, unless a policy reaches a critical fraction of minority mentors, the economy will eventually fall back to the undesirable steady state. This might explain why different empirical studies can reach opposing conclusions depending on whether a certain affirmative action policy has been "strong enough" to push the economy above this threshold.

Moreover, we show that affirmative action policies can increase the productivity through two channels: On the one hand, the average productivity of workers becomes higher when firms optimally harness the potential of both genders. On the other hand, more minority workers enter the labor force without necessarily crowding out majority workers. This latter effect is particularly large if the policy maker focuses on subsidies for minority employment, and in industries with an under-supply of workers that match the unsaturated labor market modeled here.

Finally, we compare different policies such as educational subsidies and workplace hiring quotas under varying assumptions on wage determination. We show that educational subsidies and labor market quotas are equivalent with respect to utilitarian welfare, under the condition that the competitive environment allows for type-specific wages. However, the optimal policies lead to crowding-out of majority workers, which may be an issue when selling the policy to an electorate. In that instance, subsidies offer more flexibility, and the policy maker can design them so that no majority workers are crowded out in the short term, and when the mentoring boost for the majority remains nearly constant, their job prospects are almost unharmed even in the long term. When wage setting is restricted for exogenous reasons (such as cultural norms or firm-intern politics), the crowding out of majority workers is even more severe. Indeed, in such a setup, the weight of the quota is shouldered entirely by the majority who now suffers from poor employment prospects and from *overinvestment* in education.

Related Literature. Our analysis provides a rationale for effective temporary affirmative action, which is in contrast to some previous theoretical research. Our opposing predictions stem from differing assumptions on the source of the hiring imbalance. Indeed, when taste-based discrimination is at work, affirmative action may well be a zero-sum game where the benefit to the minority is offset by a direct utility loss of the majority. Under statistical discrimination, larger minority participation need not translate into updated beliefs. Quite to the contrary: Under certain parameter values, employment quotas may actually reinforce negative stereotypes against certain groups Coate and Loury (1993). The intuition is simple: When minority employment is mandated by law, firms may have to hire minority members even if they are unskilled. This in turn may actually *reduce* their returns to education and thereby further lower equilibrium skill investment. A similar conclusion is reached when agents infer their personal success probability from their own type's employment history as in Chung (2000). In an unregulated market, observing successful people with similar characteristics sheds a positive light on one's own prospects in the labor market and hence encourages investment. However, these positive inferences disappear under temporary hiring restrictions, and agents will not be any more optimistic once the employment constraints are lifted. It is important to emphasize that these arguments rely purely on informational inferences and assume no direct productivity benefit from relatable role models.

Together, these models seem to suggest that affirmative action is futile at best, or downright harmful at worst. If individuals however receive a direct productivity boost from mentoring, we show that a more positive view is warranted.

The group complementarities at the heart of our model have been observed in the empirical literature. Dreher and Cox Jr. (1996) find that not only were female MBA graduates and students of color less likely to form mentoring partnerships with white men, but these missing relationships also had a tangible impact on later compensation. Indeed, students mentored by white men earned on average $16,840 more annually than those with mentors of other demographic profiles. In a similar vein, Ibarra (1992) analyzes the professional network within an advertising firm and finds that differential patterns of network connectivity helped men reap greater network returns than a woman in the same position. Bettinger and Long (2005) find that an increased share of female faculty positively influences course selection and major choice for female undergraduates in some (though not all) disciplines where women are historically underrepresented. Observations by Carrell et al. (2010) also seem to support the hypothesis of a measurable productivity boost from relatable role models. They show that assigning a female professor to mandatory introductory science and math classes significantly increases course grades for female students without hurting male performance, to the point of eradicating the gender gap in grades and STEM majors.

Contrary to Krishnamurthy and Edlin (2014), we assume no ex-ante differences across the two groups, which we argue is the more interesting case for long-term policy design.[1] The structure of our model is very similar to that of Athey et al. (2000), as they study the intertemporary promotion decision of long-lived firms. As in our model, they assume that senior workers offer an additive mentoring boost to junior employees, and that the size of this boost is an increasing function in the share of senior mentors of one's own type. Since their firms are forward looking, they essentially share the same objective function as our social planner in our case, albeit with limited labor demand. We believe that our two papers are complementary (and have purposely adapted our notation to match theirs closely for ease of comparison). They offer additional theoretical insight into the family of boost functions that admit certain steady states, while we frequently restrict our attention to a specific mentoring function and contrast different ways in which a social planner can guide myopic firms to a more productive equilibrium. They highlight the possibility of a "glass ceiling" in upper-level recruitment, and we show just how detrimental this can be to future

[1]By assuming that minorities have innately lower skill levels, Krishnamurthy and Edlin (2014) suggest that colleges should set *higher* admission standards for minority students in order to ensure that the same inferences are made on the ability of its graduates. The analysis fails to explain how those skill differences are created in the first place, and we believe that any policy recommendations won from such an analysis could thus be optimal only over a very short horizon.

talent recruitment by incorporating individual education decisions. By studying the case of an unsaturated labor market (which seems particularly relevant for high-skill sectors), we show how quotas affect both the intensive and the extensive margin of productivity.

2 Model

We study an overlapping generations model, where each period an infinite pool of new workers arrive. Each worker belongs to one of two groups. There is no difference across groups with respect to innate ability, which we assume to be distributed according to a decreasing *talent function* $x : [0, \infty) \to [0, \infty)$ irrespective of one's type. Specifically, we assume that there is a mass θ workers of ability at least $x(\theta)$, for all $\theta > 0$.

Each worker lives for two periods – in his youth, he decides whether or not to invest into education and in his adulthood he works in a competitive market with abundant employment opportunities. We call young workers students and grown up workers mentors. To simplify the analysis and isolate the most pertinent features, we here assume that education is discrete and the only determinant of later productivity. Real-world examples that fit this (admittedly stylized) description include any sectors where a diploma is the main hiring criterion.[2] In particular, we assume that an educated worker's productivity is equal to 1 and zero otherwise.

Mentoring is more effective between workers of the same type. This is reflected through the mentoring function

$$\mu : [0, 1] \to [0, \infty)$$

which takes as an argument the fraction of current educated employees of the student's own type. In line with Athey et al. (2000), we adopt the parameter $\phi = \phi_M \geq \frac{1}{2}$ to denote the fraction of majority mentors, and sometimes refer to $1 - \phi = \phi_m$ as the minority representation.

One micro foundation that justifies a concave mentoring boost is that every young person encounters potential mentors throughout his or her education of duration d. If the student connects with at least one potential mentor of his type, then he receives a mentoring boost of M. Assuming mentors of each type arrive according to an exponential distribution with rate ϕ_*, the resulting mentoring boost is given by $\mu(\phi_*) = M \cdot (1 - e^{-d\phi_*})$, which is more concave for larger d.[3]

[2] One may think of specialized exams such as the Bar license for lawyers or the PE license for engineers, which in and of themselves don't increase a candidate's effectiveness but are required to perform certain functions.

[3] In our numerical simulations, we use this mentoring boost function together with the normalization

Education is costly, but less so for more able individuals, or those that benefit from strong mentoring. Specifically, let $x \in [0, \infty)$ denote ability and $\mu \in [0, 1]$ be a measure for mentoring effectiveness. Both factors impact investment decisions since the cost of education $c : [0, \infty) \times [0, 1] \to [0, \infty)$ is linearly decreasing in both arguments,

$$c(x, \mu) = \bar{c} - x - \alpha\mu,$$

where the parameter $\alpha \in [0, \infty)$ measures the relative importance of mentoring versus innate skill. We ensure that total labor force never explodes and never completely disappears by requiring $\bar{c} \in \big(\alpha\mu(1), x(0) + \alpha\mu(\frac{1}{2})\big)$.

In the latter stage of life, workers enter an unsaturated labor market, with $n \geq 2$ firms of unlimited capacity. As long as the market is unregulated, competition between employers ensures that the market wage for workers equals their productivity, hence positive and equal to 1 for educated workers only. As we will see in Section 4, market regulations can affect the expected wage w_i for licensed professionals of each type $i \in \{M, m\}$ differently. Anticipating this wage, a student θ of type i invests into education if and only if $c\big(x(\theta), \mu(\phi_i)\big) \leq w_i$, pinning down the college cohort (and hence labor force) size of type i as

$$\ell(\phi_i, w_i) = \begin{cases} 0 & \text{if } c\big(x(0), \mu(\phi_i)\big) \geq w_i \\ x^{-1}\big(\bar{c} - w_i - \alpha\mu(\phi_i)\big) & \text{otherwise.} \end{cases}$$

As a result, the law of motion for the fraction of majority mentors $\phi \in [0.5, 1)$ in an unregulated market is given by

$$\phi^+(\phi) = \frac{\ell(\phi, 1)}{\ell(1 - \phi, 1) + \ell(\phi, 1)}.$$

This transition function ϕ^+ not only defines the limiting behavior of the economy, but also sheds light on the dynamics of the system. In particular, its fixed points $\phi = \phi^+(\phi)$ pins down the **steady states** of the economy. We call a steady state ϕ **stable** whenever a small perturbation in the representation of role models does not affect the long-term convergence, i.e., when there exists $\varepsilon > 0$ such that $\lim_{t\to\infty}(\phi^+)^t(\phi') = \phi$ for all $\phi' \in (\phi - \varepsilon, \phi + \varepsilon) \cap [0, 1]$.

Example. Let $x(\theta) = \lambda e^{-\lambda\theta}$ denote the ability function, where $\lambda > 0$ changes the *concentration* of talent. Since $\int_0^\infty x(\theta)d\theta \equiv 1$, total talent in the workforce is constant, but more of it is concentrated among the top candidates as λ grows.

$M = \frac{1}{1-e^{-d}}$, so that the maximal mentoring boost is always equal to 1 and any changes in its importance relative to ability is captured in the parameter α.

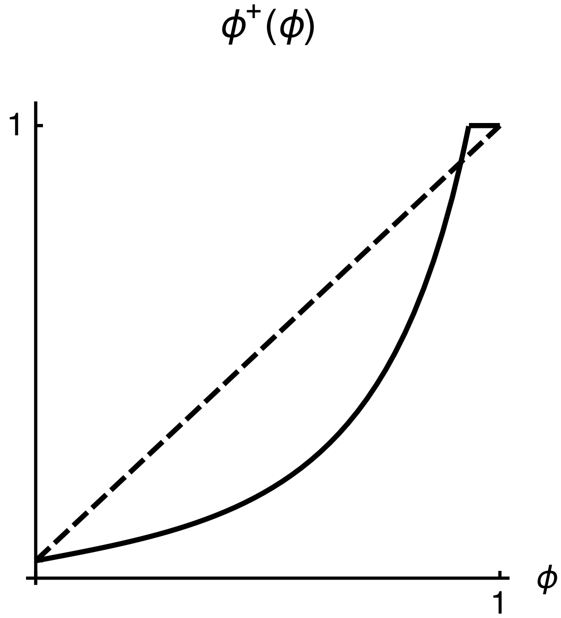


Figure 1: Transition function $\phi^{+}(\phi)$

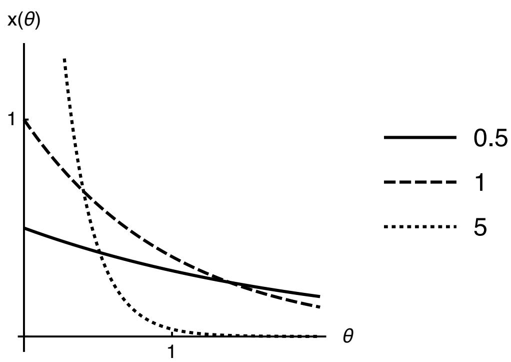
The mentoring boost is given by $\mu(\phi) = \frac{1-e^{-\phi d}}{1-e^{-d}}$, where larger levels of d increase the complementarities within each type (see Figure 2). Together with the mentoring coefficient α and the base cost $\overline{c}$, this uniquely pins down the transition function

$$\phi^{+}(\phi) = \begin{cases} 1 & \text{if } \overline{c} - \lambda - \alpha\mu(1-\phi) \geq 1 \\ \frac{\ln(\overline{c}-\lambda-\alpha\mu(\phi))-\ln\lambda}{\ln(\overline{c}-\lambda-\alpha\mu(\phi))+\ln(\overline{c}-\lambda-\alpha\mu(1-\phi))-2\ln\lambda} & \text{otherwise,} \end{cases}$$

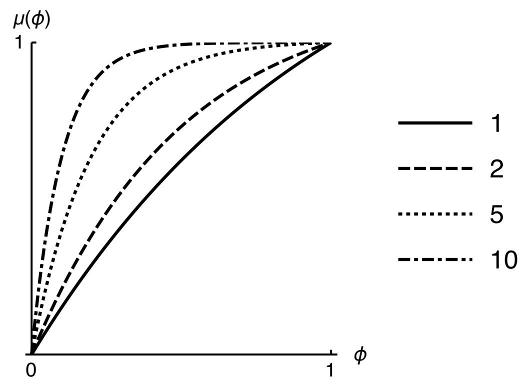
for all $\phi \in [0.5, 1]$. In particular, a completely homogeneous workforce $\phi = 1$ is a steady state if and only if talent is not too concentrated, $\lambda \leq \overline{c} - 1$, and a stable one whenever the inequality is strict. Figure 1 illustrates the transition function for $s = .15, \lambda = 1, \alpha = .5, \overline{c} = 2.1$.

A perfectly balanced workforce $\phi = 0.5$ is always a steady state by symmetry, and a stable one when talent is *sufficiently* concentrated, $\ln\lambda \geq \ln\left(\overline{c} - 1 - \alpha\mu(\frac{1}{2})\right) + \frac{1}{2}\frac{\alpha\mu'(\frac{1}{2})}{\overline{c}-1-\alpha\mu(\frac{1}{2})}$. Figure 2(c) graphically represents the stable steady states as a function of d and λ. ♦

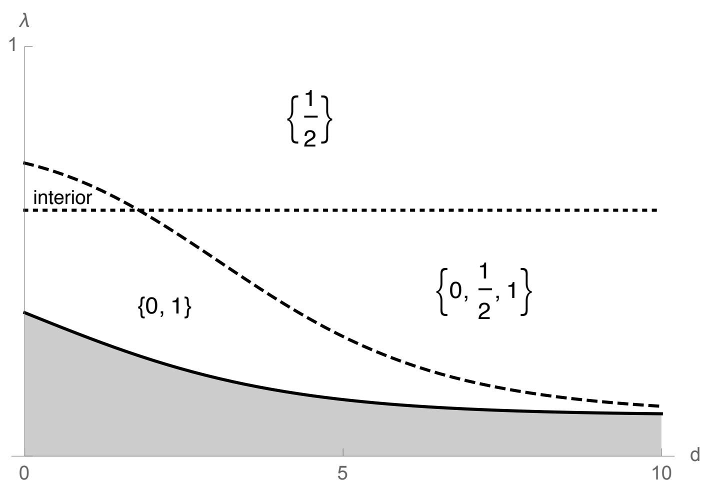
Alternative Interpretation: Job Appeal. The additive productivity boost through mentoring is just one interpretation of our model. Although supported by empirical evidence, the salience of this feature may vary across sectors. To underline the relevance of our model, we here discuss an entirely different setup without mentoring tradeoffs that nonetheless



(a) ability function



(b) mentoring boost



(c) Stable steady states as a function of complementarity factor d and talent concentration λ, for $\alpha = 0.5$ and $c = 1.6$. The shaded area corresponds to invalid parameter values where the workforce completely disappears at $\phi = 0.5$.

Figure 2: Parametrized version of the model.

yields the same dynamics.

In particular, suppose that the *appeal* of a particular profession depends on the race and gender representation of its current workforce. For instance, recent political events have drawn criticism about the underrepresentation of African Americans in the police force. It is often argued that the racial imbalance in law enforcement is both cause and result of poor trust ratings among minorities.[4] In a similar way, young girls may steer away from predominantly male fields not because they doubt their own chances of success (as in Chung (2000)), but because they inherently enjoy having professors and colleagues of the same gender.[5] By interpreting $\mu(\cdot)$ as extra utility enjoyed from working in the high-skilled industry, such a setup is equivalent to the previous model.

3 Dynamics and Welfare

3.1 Steady States

As already noted, an unregulated market, will always converge to a steady state in the long run. Due to symmetry, it is immediate that a **balanced labor force** ($\phi = \frac{1}{2}$) is always a steady state.

Proposition 1 *(i) $\phi = \frac{1}{2}$ is always a steady state. It is stable if and only if*

$$\alpha\mu'\left(\frac{1}{2}\right) < \underbrace{2\ell\left(\frac{1}{2}, 1\right)}_{\textit{labor force}} \underbrace{\left(-x'\left(\ell\left(\frac{1}{2}\right)\right)\right)}_{\textit{marginal loss in ability}}$$

(ii) $\phi \in \{0, 1\}$ is a stable steady state if and only if

$$x(0) < \bar{c} - 1.$$

Proof. (i) The balanced labor force is a stable steady state if $(\phi^+)'(0.5) < 1$ since then, there exists an $\varepsilon > 0$ such that in any open set $(0.5, 0.5 + \varepsilon')$ with $\varepsilon' < \varepsilon$, $0.5 < \phi^+(\phi) < \phi$.

[4] In a recent article published in the *Wall Street Journal*, Kesling and McWhirter (2015) mention the struggle to recruit and retain African-Americans, which they argue is at least partly explained by "black distrust of police departments". Diversity at individual departments may vary specifically *because* high-level recruiters seek to counteract this imbalance. Indeed, citing the recruitment officer for St.Louis County Police Department, they write that good quality minority applicants "can write their own ticket" by filling out applications for multiple departments and then picking the most appealing one. This preferential treatment of *qualified* minority applicants under Affirmative Action efforts gives further credence to the type-dependent wage assumption in Section 4.2.

[5] The effects measured in Bettinger and Long (2005) may in fact measure either expected mentoring benefits or pure taste preference.

Thus, for any $\phi \in (0.5, 0.5+\varepsilon)$, $\lim_{t\to}(\phi^+)^t(\phi)$ exists. Whenever $\phi' = \lim_{t\to}(\phi^+)^t(\phi) > 0.5$, it must be that $\phi^+(\phi') = \phi'$ which leads to a contradiction.

(ii) If $x(0) < \bar{c} - 1$, then there exists an ε such that for $\phi \in (1-\varepsilon, 1)$ no minority student invests in investment. Thus, $\phi^+(\phi) = 1$. ■

This proposition provides conditions under which both completely homogeneous and balanced steady states can be stable at the same time. A balanced labor force is stable if the ability and labor force drop caused by putting the labor force out of balance (i.e., increasing ϕ) dominates the marginal benefit of better mentoring boost. Moreover, if for the most talented student $x(0)$ investment is not optimal absent any mentoring, then a fully homogeneous steady state is also stable. Both conditions are local and thus, there is a plethora of mentoring and ability functions that satisfy both conditions. This hints to the possible need of policies if the economy happens to be in a sub-optimal steady state.

3.2 Welfare

Let the surplus generated by a generation with z_m minority and z_M majority students investing be given by

$$\begin{aligned}\pi(z_m, z_M, \phi) &= \int_0^{z_m} x(\theta) + \alpha\mu(1-\phi) + 1 - \bar{c}d\theta \\ &+ \int_0^{z_M} x(\theta) + \alpha\mu(\phi) + 1 - \bar{c}d\theta.\end{aligned}$$

A social planner, who can fully regulate the number of students who invest, solves

$$V(\phi;\delta) = \max_{z_m, z_M} \pi(z_m, z_M, \phi) + \delta \cdot V\left(\frac{z_M}{z_m + z_M}; \delta\right).$$

Let the solution to this problem be $\ell_m^*(\phi;\delta), \ell_M^*(\phi;\delta)$. Thus, $\ell_m(\phi,1) = \ell_m^*(\phi;0)$, $\ell_M(\phi,1) = \ell_M^*(\phi;0)$.

Proposition 2 *Whenever $\phi^+(\phi)\mu'(\phi) < (1-\phi^+(\phi))\mu'(1-\phi)$ for all $\phi > 0.5$, then it is optimal to hire some minority workers who do not produce surplus and to exclude some majority workers who would produce surplus for any $\phi > 0.5$ and $\delta > 0$. Analogously, if the inequality is reversed, then it is optimal to hire some majority workers who do not produce surplus and to exclude some minority workers who would produce surplus for any $\phi > 0.5$.*

Proof. The derivatives of the objective function in the maximization problem of V at $\ell_m(\phi,1), \ell_M(\phi,1)$ is

$$\delta \cdot V_\phi\left(\phi^+(\phi);\delta\right) \cdot \frac{-\ell_M(\phi,1)}{(\ell_m(\phi,1)+\ell_M(\phi,1))^2} \quad , \quad \delta \cdot V_\phi\left(\phi^+(\phi);\delta\right) \cdot \frac{\ell_m(\phi,1)}{(\ell_m(\phi,1)+\ell_M(\phi,1))^2}.$$

Thus, it remains to show that $V(\cdot;\delta)$ is decreasing which can be shown analogously to the proof of Proposition 1 in Athey et al. (2000). ■

This proposition is analogous to Proposition 1 in Athey et al. (2000). It is difficult to make stronger statements as the powerful tools of dynamic programming only apply for concave and monotonic π. Nevertheless, we can say the following.[6]

Proposition 3 *(i) Whenever* $V(0.5;0) = \max_\phi V(\phi;0)$*, for every* ϕ $\lim_{\delta\to 1}(1-\delta)V(\phi;\delta) \to V(0.5;\delta)$*.*

(ii) Whenever $V(0.5;0) = \max_\phi V(\phi;0)$*, the first-best policy in any state* $\phi \in (0.5, 1]$ *is such that* $\ell^*_M(\phi,\delta) > \ell^*_m(\phi,\delta)$*.*

(iii) Whenever $\frac{\ell^*_M(\phi;\delta)}{\ell^*_M(\phi;\delta)+\ell^*_m(\phi;\delta)} < \phi^+(\phi)$*,* $\ell^*_M(\phi;\delta) < \ell_M(\phi,1)$ *and* $\ell^*_m(\phi;\delta) > \ell_M(\phi,1)$*, i.e. some qualified majority workers remain unemployed and some unqualified minority workers get employed.*

Proof. (i) Note that $V(0.5;\delta) = \frac{1}{1-\delta}V(0.5;0)$. Thus, $(1-\delta)\left(\pi\left(z,z,\phi\right) + \delta V(0.5;\delta)\right) \to_{\delta\to 1}= V(0.5;0)$ for all z.

(ii) If the policy jumps straight to $\phi = \frac{1}{2}$, i.e., $z_m = z_M = z$, then a marginal change in z_m leads to

$$\frac{\partial \pi}{\partial z_m}(z,z) + \delta \underbrace{V_\phi(0.5,\delta)}_{=0} \cdot \frac{-z}{(2z)^2} < 0,$$

i.e., decreasing z_m increases profits. Similarly, a marginal increase in z_M increases profits if there are some majority workers who do not work despite being sufficiently able. Thus, the optimal policy should never jump fully to the perfectly balanced steady state.

(iii) Let us assume that the optimal policy dictates $\frac{\ell^*_M(\phi;\delta)}{\ell^*_M(\phi;\delta)+\ell^*_m(\phi;\delta)} \equiv \tilde{\phi}^+ < \phi^+(\phi)$. In that case, either $x(\ell^*_M(\phi;\delta)) + \alpha\mu(\phi) + 1 - \overline{c} > 0$ or $x(\ell^*_m(\phi;\delta)) + \alpha\mu(1-\phi) + 1 - \overline{c} < 0$. Denote $L^* \equiv \ell^*_M(\phi;\delta) + \ell^*_m(\phi;\delta)$. Let us first assume that $x(L^*(1-\tilde{\phi}^+)) + \alpha\mu(1-\phi) + 1 - \overline{c} < 0$ and $x(L^*\tilde{\phi}^+) + \alpha\mu(\phi) + 1 - \overline{c} \le 0$. Then,

$$\begin{aligned} \frac{\partial}{\partial L^*}\pi(L^*(1-\tilde{\phi}^+), L^*\tilde{\phi}^+, \phi) \quad &= \quad (1-\tilde{\phi}^+)(x(L^*(1-\tilde{\phi}^+)) + \alpha\mu(1-\phi) + 1 - \overline{c}) + \\ &\quad\quad \tilde{\phi}^+(x(L^*\tilde{\phi}^+) + \alpha\mu(\phi) + 1 - \overline{c}) < 0. \end{aligned}$$

Thus, it is a profitable deviation to decrease the total labor force, holding $\tilde{\phi}^+$ constant. Thus, $x(L^*\tilde{\phi}^+) + \mu(\phi) + 1 - \overline{c} > 0$. Analogously, ie we assume $x(L^*\tilde{\phi}^+) + \mu(\phi) + 1 - \overline{c} > 0$, we can show that $x(L^*(1-\tilde{\phi}^+)) + \mu(1-\phi) + 1 - \overline{c} < 0$. Thus, the sufficiently qualified

[6] Note that the existence of a solution can be shown with standard arguments using the contraction mapping theorem.

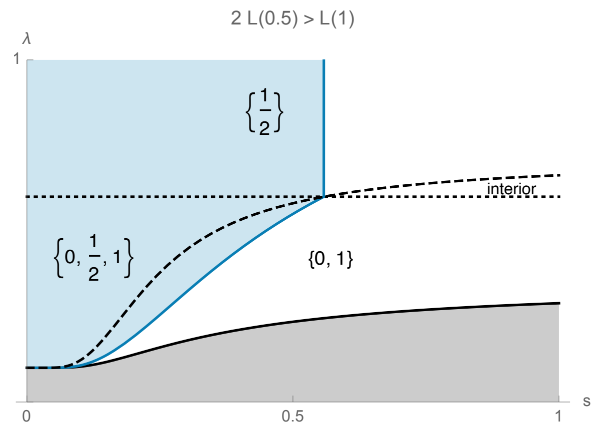


Figure 3: Higher labor force in balanced economy

majority workers must not work and unqualified minority workers must work under the optimal policy.. ■

This proposition shows that even if the goal is to reach equal labor force across groups, affirmative action must approach the goal slowly. If the highest myopic outcome is to have a balanced work force, then it is optimal to eventually move to a balanced work force as long as the social planner is sufficiently patient.

Example. In the example introduced above, the workforce is *larger* under balanced mentoring than in a homogeneous setting if and only if talent is sufficiently concentrated, $\ln \lambda \geq 2\ln(\bar{c} - 1 - \alpha\mu(\frac{1}{2})) - \ln(\bar{c} - 1 - \alpha)$. Figure 3 highlights this region.

♦

4 Affirmative Action Policies

Having thus shown that the policy maker can generally improve welfare through market intervention, we here turn our focus to the practical implementation of such a policy. Indeed, since educational decisions are taken individually, a social planner can merely nudge the independent actors into the right direction. In other words, showing that long-term welfare is higher under different investment decisions doesn't mean that they *can be* modified in such a way. However, as we show in this section, the policy maker has two options to implement desired investment levels: Direct modification in the cost of education (through a tuition raise and fellowships), as well as labor market regulations.

However, before we define the specific policies, we add slightly more notation in order to simplify the discussion. First, we call a quota $\hat{\phi}$ **binding** at diversity level ϕ if it requires a more diverse workforce than what would myopically be implemented, $\phi^+(\phi) > \hat{\phi}$. Moreover, while an unsaturated labor market guarantees employment security for all students, this may no longer be the case in a regulated market. To clarify this effect, we thus use (M, m) and (ℓ_M, ℓ_m) to denote the mass of **educated** and **employed** individuals of each type, respectively.

4.1 Educational incentives

The first, most direct market intervention directly modifies the cost-benefit analysis of prospective students. Indeed, by implementing fellowship or tuition hikes Δc_i such that

$$\Delta c_i = \overline{c} - 1 - x(\ell_i) - \alpha\mu(\phi_i),$$

the policy maker can implement any desired labor force participation ℓ_i.[7] The flexibility of this approach ensures that any optimal market intervention (1) can be implemented this way.

By Proposition 3, socially optimal intervention always includes crowding out among majority workers if it is optimal to increase the fraction of minority workers. As we briefly alluded to in the introduction, this may make it harder for the government to sell such a strategy to the electorate. In such a case, educational incentives have the added benefit of allowing for only one-sided adjustments, whereby *only* the minority tuition is lowered (e.g. through fellowships). If the mentoring function is sufficiently concave and the majority mentor boost is only minimally affected by the move towards a more balanced workforce, this offers a politically palatable way of pushing the economy towards a higher-welfare state.

One critical observation from our model is that we here discuss fellowships $\Delta c_m > 0$ that are available to *all* interested minority students, rather than selective fellowships for only the most promising candidates. If fellowships are only available to the most talented individuals of each group, it is straightforward to see that such a fellowship will have *no effect* on the extensive margin unless the potential candidate pool exceeds $\ell(\phi, 1)$, that is the measure of top minority candidates who would also otherwise invest into education.[8] This may explain why studies such as Prenovitz et al. (2015) fail to observe additional minority recruitment for competitive scholarship programs with a very limited budget.

[7] If budget balance is a concern, note that the policy maker can always recoup any imbalance through a general tax levied on all agents irrespective of type or education status.

[8] Total demand for a fellowship of size Δc_m is given by $\ell(\phi, 1 + \Delta c_m)$, so any larger pool will not have an effect either.

4.2 Labor Force Quotas

Alternatively, the policy maker can restrict the hiring decisions of firms by requiring them to employ *at least*[9] a fraction $(1-\hat{\phi}) \in [0, \frac{1}{2}]$ of *qualified* minority workers.[10] As firm competition ultimately decides on market wage and the size of the labor force, the outcomes of such a quota depend on whether the market allows for wage differentials based on minority membership.

Type-specific wage. In a fully unregulated market where wage is determined *only*through market forces, any imbalance in the supply of educated workers, $(1-\hat{\phi})M > \hat{\phi}m$, will affect the market wage for majority and minority workers. In such a situation, educated minority workers are in short supply and firms pay a premium in order to attract them, while the oversupply of majority workers forces their wage down to zero.

Adopting a general equilibrium approach, we write down the market-clearing equations for educated labor under a binding quota $\hat{\phi} < \phi^+(\phi)$ as

$$\begin{cases} 0 & = 1 - \hat{\phi} w_M - (1-\hat{\phi}) w_m, \\ w_M & = c\big(x(\hat{\phi}L), \mu(\phi)\big), \\ w_m & = c\big(x((1-\hat{\phi})L), \mu(1-\phi)\big). \end{cases}$$

The first equation is the zero-profit condition for competitive firms, stating that the marginal benefit of hiring a high-productivity worker (at the required diversity level) is equal to zero. The second and the last equation stem from workers' individual rationality constraints: At the given wage level, the marginal worker needs to be exactly indifferent about investing in education. Taken together, these market clearing conditions uniquely determine the size of the total labor force L as well as type-specific wages $w_M < 1 < w_m$.[11]

Imposing such a quota is less flexible than educational incentives as it delegates the decision over labor force size to myopic firms. Nevertheless, since only the *composition* of the work-force affects future mentoring, while the *size* impacts only short-term welfare, the following result shows that there is no efficiency loss in doing so.

[9]For ease of exposition, we do not consider quotas that set an upper bound on the minority representation, but the same analysis would go through, albeit with a higher expected wage for *majority* candidates.

[10]If a quota does not contain restrictions on qualification, firms could always costlessly meet any quota by hiring unqualified minority workers at a wage of zero, and thus be completely ineffective.

[11]The bounds on individual wage are obtained since the quota is *binding*, which directly rules out $(w_M, w_m) = (1, 1)$. By the zero-profit condition, one wage is thus bigger and one smaller than one. And since $w_M > 1 > w_m$ would only attract *even more* majority employees, expected equilibrium wages have to satisfy the bounds outlined above.

Proposition 4 *Whenever it is optimal to incentivize a more* diverse *workforce,*[12] *the optimal size* $L^* = \ell^*_m + \ell^*_M$ *is equal to the market equilibrium under type-specific wage.*

Proof. A simple change of variables in Section 3.2 yields

$$\max_{L,\hat{\phi}} \pi\big(\hat{\phi}L, (1-\hat{\phi})L, \phi\big) + \delta V(\hat{\phi}, \delta).$$

Since L only affects *current* welfare, the first order condition implies

$$0 = \phi^* \frac{\partial \pi}{\partial \ell_M} + (1-\phi^*)\frac{\partial \pi}{\partial \ell_m} = \phi^*\big(1 - c\big(x(\phi^* L), \mu(\phi)\big)\big) + (1-\phi^*)\big(1 - c(x((1-\phi^*)L), \mu(1-\phi))\big),$$

which corresponds exactly to the market-clearing equation above. ■

Common wage. In some industries, social or legal pressure prohibits paying unequal wage to employees in the same position. In this case, the zero-profit condition uniquely pins down $w = 1$.[13] While all m educated minority workers find employment, firms only hire $\frac{\hat{\phi}}{1-\hat{\phi}}m$ majority workers under a binding quota. This is factored into workers cost-benefit analysis as they contemplate education, hence minority investment *and* education are unaffected and intervention is felt exclusively by the majority.

Proposition 5 *When type-specific wages are infeasible, labor force quotas do affect neither minority investment nor recruitment. Majority investment and employment are reduced, but the latter more so than the former, leading to over-investment in education.*

Proof. Since the minority is guaranteed certain employment at wage 1, the relevant expected wage is equal to the unregulated case discussed in Section 2, implying $m = \ell(1-\phi, 1)$. The partial employment of majority workers however reduces the expected wage to $\frac{\hat{\phi}}{1-\hat{\phi}}\frac{m}{M} < 1$ under a binding quota. The indifference condition

$$c\big(x(M), \mu(\phi)\big) = \frac{\hat{\phi}}{1-\hat{\phi}}\frac{\ell(1-\phi, 1)}{M}$$

uniquely pins down majority investment M. Indeed, note that the left side is strictly increasing in M and the right side is strictly decreasing in M, with $1 = c\big(x(\ell(\phi, 1)), \mu(\phi)\big) >$

[12] The result can be adapted to situations with labor market quotas that bind majority employment from *below*, but to simplify the discussion we do not formally define such quotas here.

[13] Indeed, whenever $w < 1$, a firm can attract all desired workers with a slightly higher wage $w + \varepsilon$. Hiring m educated minority and $\min\{1, \frac{\hat{\phi}}{1-\hat{\phi}}\}$ majority workers represents a profitable deviation for at least one firm. Similarly, when $w > 1$, any active firm makes negative profit – but if no firm is active, workers would be willing to work for any positive wage $w > 0$ and thus create a business opportunity.

$\frac{\hat{\phi}}{1-\hat{\phi}} \frac{\ell(1-\phi,1)}{\ell(\phi,1)}$ since the quota binds. This implies that educational investment is reduced, $M < \ell(\phi, 1)$, and that expected wage is strictly lower than one, implying imperfect employment $M > \ell_M$. ■

The partial employment creates an additional negative externality on the majority, which greatly reduce the appeal of workplace quotas in situations where wage is sticky or subject to social scrutiny. Even so, real-world firms have many different ways of implementing favorable work conditions for a minority that don't rely on wage premiums. An example are workplace or schedule accommodations that particularly appeal to the target minority. The reality thus likely lies in between the two situations outlined here, making workforce quotas an effective, but less versatile and politically more challenging policy tool.

5 Conclusion

By analyzing the far-reaching impact of workplace complementarities on long-term welfare, we hope to contribute to the discussion on affirmative action by reconciling some arguments from both sides. For instance, opponents of positive discrimination often criticize the *persistence* of such policies. The very fact that lower admission standards for minority are *still* viewed as necessary after decades, the argument goes, demonstrates their ineffectiveness in bringing about lasting change.[14] Our analysis highlights the shortcomings of such an inference: Indeed, there is room for welfare improvement through policy design in many situations, and by carefully weighing the present cost against future efficiency gains, the *optimal* policy implements such a change over infinitely many periods. Rather than rendering itself obsolete after a short period of time, an optimal intervention nudges the economy slightly closer to the optimal steady state in every period. Certainly, these nudges get smaller and less costly as inequalities shrink, but they never fully stop.

That being said, minority support programs *can* be completely ineffective at minority recruitment if they are chosen 'too small'. As we highlight in Section 4, merit-based minority scholarships must be available in big enough quantity in order to affect the extensive margin of minority education.

References

Athey, Susan, Christopher Avery, and Peter Zemsky, "Mentoring and Diversity," *American Economic Review*, 2000, *90* (4), 765–786.

[14]See for instance alumni.stanford.edu/get/page/magazine/article/?article_id=43448, retrieved 08/18/15.

Bettinger, Eric P and Bridget Terry Long, "Do faculty serve as role models? The impact of instructor gender on female students," *American Economic Review*, 2005, pp. 152–157.

Carrell, Scott E, Marianne E Page, and James E West, "Sex and science: How professor gender perpetuates the gender gap," *The Quarterly Journal of Economics*, 2010, *125* (3), 1101–1144.

Chung, Kim-Sau, "Role models and arguments for affirmative action," *American Economic Review*, 2000, pp. 640–648.

Coate, Stephen and Glenn C Loury, "Will affirmative-action policies eliminate negative stereotypes?," *The American Economic Review*, 1993, pp. 1220–1240.

Dreher, George F. and Taylor H. Cox Jr., "Race, Gender, and Opportunity: A Study of Compensation Attainment and the Establishment of Mentoring Relationships.," *Journal of Applied Psychology*, 1996, *81* (3), 297 – 308.

Ellison, Glenn and Ashley Swanson, "The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competitions," Technical Report, National Bureau of Economic Research 2009.

Fang, Hanming and Andrea Moro, "Theories of Statistical Discrimination and Affirmative Action: A Survey," in Jess Benhabib, Matthew O. Jackson, and Alberto Bisin, eds., *Handbook of Social Economics, Vol. 1A*, The Netherlands: North-Holland, 2011, pp. 133–200.

Ibarra, Herminia, "Homophily and Differential Returns: Sex Differences in Network Structure and Access in an Advertising Firm," *Administrative Science Quarterly*, 1992, *37* (3), pp. 422–447.

Kesling, Ben and Cameron McWhirter, "Percentage of African-Americans in U.S. Police Departments remains flat since 2007," *The Wall Street Journal*, May 14 2015.

Krishnamurthy, Prasad and Aaron Edlin, "Affirmative Action and Stereotypes in Higher Education Admissions," Working Paper 20629, National Bureau of Economic Research October 2014.

Milkman, Katherine L., Modupe Akinola, and Dolly Chugh, "What Happens Before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations," April 2014. unpublished.

Prenovitz, Sarah J., Gary R. Cohen, Ronald G. Ehrenberg, and George H. Jakubson, "An Evaluation of The Mellon Mays Undergraduate Fellowship's Effect on PhD Production at Non-UNCF Institutions," Working Paper 21451, National Bureau of Economic Research August 2015.

Sotomayor, Sonia, *Schuette v. Coalition to Defend Affirmative Action. 572 U.S.* 2014.