Mentoring
and the Dynamics of Affirmative Action

Michèle Müller-Itten (University of Notre Dame) & Aniko Öry (Yale SOM)

Yale School of Management

Penn State Theory Seminar

November 20, 2020

Motivation

1. **Mentor relationships are stronger between members of the same sociodemographic group (typically race or gender)**
   Dreher and Cox Jr. (1996), Ibarra (1992)
2. **Lack of role models affects academic performance and labor market outcomes of minority students (not explained by innate ability)**
   Bettinger and Long (2005), Card and Giuliano (2016), Carrell et al (2010),
   Dee (2004, 2005, 2007), Ellison and Swanson (2009), Fairlie et al. (2014)
3. **Achievement differences arise early and lead to different educational choices**
   Bayer and Rouse (2016)

$\Rightarrow$ If labor force evolution is governed by education choices, then dynamic reinforcement may exacerbate workforce imbalance.

Research Question

Key trade-off:

$$\begin{array}{ccc} \text{Mentoring efficiency} & \Leftrightarrow & \text{Use of innate talent} \\ \downarrow & & \downarrow \\ \text{mentor capacity} & & \text{talent dispersion} \end{array}$$

Source of inefficiency: Juniors do not internalize their positive mentoring externality when making education decisions.

Research Questions:

1. **Positive:** What are steady state workforce compositions?
2. **Normative:**
   - What is the optimal steady state workforce composition?
   - What is the optimal workforce composition?
3. **Policy implication:** What are the implications for the debate surrounding Affirmative Action (AA) policies?
   $\rightarrow$ We define a workforce to be **fair** when two people with the same talent make the same education choice.

Key observations

1. **Positive:** In any steady state, either the majority is over-represented or the minority has a representation of greater than 50%.
2. **Normative:**
   - If the mentor capacity is high or the talent dispersion is high, the welfare-optimal steady state is mixed and close to **fair**.
   - If the mentoring capacity is sufficiently large, the minority workers are **over**-represented in the welfare-maximizing labor force
3. The optimal intervention in such a situation is persistent.

Motivating Example

- Overlapping unit-mass generations, each generation lives for two periods.
- **Juniors** invest into costly education.
- **Seniors** produce a surplus of 1 if educated and **10 role models** for the future generation are drawn at random.
- In each generation, 80% belong to the **majority** group $(i = 1)$ and 20% to the **minority** group $(i = 2)$.
- **Talent** is distributed equally in both groups (and private info):
  - $\frac{1}{4}$ has high talent $H \rightarrow$ no cost of education
  - $\frac{3}{4}$ low talent $L \rightarrow$ positive cost $c > 0$ of education
- Each junior receives a payoff of 1 if one of the role models is from the own group, but the realization of role models is only known after the education decision.
- A L-type junior invests if and only if $c < 1 - (1 - \phi)^{10}$, where ϕ is the fraction of own group seniors.

Motivating Example

- Set c such that a L-type junior invests if and only if $\phi > 0.35$.
- (i) is the only steady state where all investment decisions are individually rational.
- (ii) and (iii) might seem more "fair"?
- (iv) maximizes welfare.

Composition	(i)		(ii)		(iii)		(iv)	
	H	L	H	L	H	L	H	L
Majority participation	•	•	•		•	•	•	
Minority participation	•		•		•	•	•	•
% majority workers	94%		80%		80%		50%	
Total surplus	1.92		1.98		1.95		2.01	

Literature

- **Taste-based discrimination:** Becker (1957) - zero sum game
- **Statistical discrimination:** Coate and Loury (1993), Fang and Moror (2011), Fryer Jr (2007)
- **Biased beliefs:** Boren et al. (2018), Bordalo et al. (2019)
- **Learning about success chance from role models:** Chung (2000)
- **Mentoring and optimal promotion decisions:** Athey et al. (2000)

Model

- Time: $t = 0, 1, 2, \ldots$
- Overlapping unit mass generations: each agent lives for at most two periods (junior $\rightarrow$ senior)
- Each agent is indexed by
  - **talent** $x \in \mathbb{R}$, $x \sim F$
  - **group membership** i: $\begin{cases} \text{majority } (i=1) \text{ with probability } b \geq 0.5 \\ \text{minority } (i=2) \text{ with probability } 1-b \leq 0.5 \end{cases}$
- $b =$ **majority share**
- Upon birth, a junior decides whether to participate in the labor force to become a senior and earn a wage of w.
- Each senior produces a surplus of $\pi \Rightarrow$ with free entry of firms, $w = \pi$
- Education is costly:

$$c - \underbrace{x}_{\text{talent}} - \underbrace{\tilde{\mu}(L_i, L_{\neg i}, \ell_i, \ell_{\neg i})}_{\text{mentoring boost}}$$

  where $\mathbf{L} = (L_1, L_2)$ represent the mass of seniors of groups $1, 2$ and $\ell = (\ell_1, \ell_2)$ the participating juniors of groups 1 and 2.

Junior's IR constraint

- Given a senior workforce of $\mathbf{L} = (L_1, L_2)$ and wages w_1, w_2, and given other juniors invest so that the junior workforce is $\boldsymbol{\ell} = (\ell_1, \ell_2)$, then it is individually rational for a junior of talent x to invest if and only if

$$c - x - \tilde{\mu}\,(L_i, L_{\neg i}, \ell_i, \ell_{\neg i}) \leq w_i.$$

- We say the junior workforce $\ell = (\ell_1, \ell_2)$ is individually rational if

$$\begin{cases} \ell_1 = \qquad\quad b\,(1 - F(c - \tilde{\mu}(L_1, L_2, \ell_1, \ell_2) - w_1)) \\ \ell_2 = (1-b)\,(1 - F(c - \tilde{\mu}(L_2, L_1, \ell_2, \ell_1) - w_2)) \end{cases}$$

  where $w_1 = w_2 = \pi$ in an unregulated economy with free entry.

- Let us denote the senior workforce in period t by $\mathbf{L}^t$ and the junior workforce in period t by $\boldsymbol{\ell}^t$.
  $\rightarrow$ The dynamic system of labor force participation is characterized by

$$\mathbf{L}^{t+1} = \boldsymbol{\ell}^t$$

Constant labor force and Steady State

- We are primarily interested in the group representation of a **constant labor force**, where $\boldsymbol{L}^t \equiv (\phi L, (1-\phi)L)$ where we call
  - $\phi \in [0,1]$ the labor market **composition** and
  - $L \in [0,1]$ its total **size**.
- L is a **steady state** of the dynamic system if for $\mathbf{L}^0 = L$, $\mathbf{L}^t = L$ for all t.
- L is **(Lyapunov) stable** if for all $\epsilon > 0$, there exists a $\delta > 0$ such that if $\|\boldsymbol{L}^0 - \hat{\boldsymbol{L}}\| < \delta$, then $\|\boldsymbol{L}^t - \hat{\boldsymbol{L}}\| < \epsilon$ for all $t > 0$.

Mentorship function

We impose the folowing assumptiosn on $\tilde{\mu}(L_i, L_{\neg i}, \ell_i, \ell_{\neg i})$:

1. **Homogeneity of degree zero**: $\tilde{\mu}(L_i, L_{\neg i}, \ell_i, \ell_{\neg i}) \equiv \tilde{\mu}(kL_i, kL_{\neg i}, k\ell_i, k\ell_{\neg i}) \qquad \forall k > 0.$
2. **Derivatives:** $\tilde{\mu}$ is continuously differentiable, and
$$\frac{\partial \tilde{\mu}}{\partial \ell_i} \leq 0, \quad \frac{\partial \tilde{\mu}}{\partial \ell_{\neg i}} \leq 0, \quad \frac{\partial \tilde{\mu}}{\partial L_i} > 0 \quad \text{over } (0,1]^4.$$
($\partial \tilde{\mu}/\partial L_{\neg i} > 0$, if cross-group mentorship is effective. $\partial \tilde{\mu}/\partial L_{\neg i} < 0$ if search frictions make it harder to find an own-group mentor in a senior workforce that is dominated by the opposite group.)
3. **Adding an own-group senior is weakly more beneficial than adding an opposite-group senior**:
$$\frac{\partial \tilde{\mu}}{\partial L_i} \geq \frac{\partial \tilde{\mu}}{\partial L_{\neg i}},$$
4. **Adding an own-group junior lowers mentoring weakly more than an opposite-group junior**:
$$\frac{\partial \tilde{\mu}}{\partial \ell_i} \leq \frac{\partial \tilde{\mu}}{\partial \ell_{\neg i}}.$$

Welfare

Let $\mu(\phi) := \tilde{\mu}(\phi, 1-\phi, \phi, 1-\phi)$.

We denote the total surplus of a constant labor force of composition ϕ and size L by

$$S(\phi, L) = \quad b \int_{x \geq \hat{x}_1} \big(\pi - c + x + \mu(\phi)\big) dF(x)$$

$$+ (1-b) \int_{x \geq \hat{x}_2} \big(\pi - c + x + \mu(1-\phi)\big) dF(x)$$

where $\hat{x}_1 = F^{-1}\Big(1 - \frac{\phi}{b} L\Big)$ and $\hat{x}_2 = F^{-1}\Big(1 - \frac{1-\phi}{1-b} L\Big)$ denote the marginal talent of group 1 and 2 workers, respectively.

Fairness

- We say a labor market of constant workforce $(\phi L, (1-\phi)L)$ is **fair** if no individual could be made better off by being born into the other group, i.e. $w_1 + \mu(\phi) = w_2 + \mu(1-\phi)$.
- We say that a labor force is **dominated by the majority (minority)** if more than half of the labor force belongs to that group, $\phi > (<)\, 0.5$.
- We say that a labor force is **over-represents the majority (minority)** when the share of workers belonging to that group is larger than the corresponding population share, $\phi > (<)\, b$.

Note that whenever $b > 0.5$, a labor force can be dominated by the majority yet still over-represent the minority.

Mentor capacity

$\tilde{\mu}$ is parameterized by the **mentoring capacity** $q > 0$ so that

- $\lim_{q \to 0} \mu_q(\phi) = 0, \quad \lim_{q \to \infty} \mu_q(\phi) = 1,$ and $\lim_{q \to \infty} \mu_q'(\phi) = 0.$
- Letting $M_q(\phi) := \phi\mu_q(\phi) + (1-\phi)\mu_q(1-\phi)$ denote the **total surplus generated by mentorship**, we assume that for any $\phi > 0.5$, there exists $Q_\phi \in$ such that

$$M_q'(\phi) < 0 \qquad \forall q \geq Q_\phi.$$

- For any $\delta > 0$, there exists a bound $K_\delta > 0$ such that

$$\|\nabla \tilde{\mu}_q(\phi, 1-\phi, \phi, 1-\phi)\|_\infty < K_\delta \mu_q'(\phi) \qquad \forall \phi \in (\delta, 1-\delta),\ \forall q > 0.$$

Mentor capacity parametrization

1) **Mentoring and matching frictions:** $n(L_1 + L_2)$ seniors are matched to $n(\ell_1 + \ell_2)$ juniors randomly where members of the same group i are linked with probability p_{ii} and members of opposite groups with probability $p_{ij} < p_{ii}$:

$$p_{ij} = \begin{cases} \frac{sq}{n\ell_i} + \frac{(1-s)q}{n(\ell_1+\ell_2)} & \text{if } i = j \\ \delta\frac{(1-s)q}{n(\ell_1+\ell_2)} & \text{otherwise.} \end{cases}$$

- On average each mentor is matched with q mentees;
- A fraction $s \in [0, 1)$ is drawn same-groups juniors;
- An acrooss-group mentor assignment is effective only with probability δ.

As n grows this converges to:

$$\tilde{\mu}(L_i, L_{\neg i}, \ell_i, \ell_{\neg i}) = 1 - e^{-sq\frac{L_i}{\ell_i} - (1-s)q\frac{L_i + \delta L_j}{\ell_1+\ell_2}}.$$

Mentor capacity parametrization

2) **Role Models:**
   - In every generation, q role models are randomly appointed.
   - A junior with $k \in \mathbb{N}_0$ same-group role models enjoys a mentorship boost of $1 - \delta^k$ for some $\delta \in [0, 1)$.
   - The expected mentorship boost to a group-i junior is then equal to

$$\tilde{\mu}(L_i, L_{\neg i}, l_i, l_{\neg i}) = 1 - \left(\frac{\delta L_i + L_{\neg i}}{L_i + L_{\neg i}} \right)^q .$$

Talent dispersion

F with support $(\underline{x}_F, \bar{x}_F)$ is parametrized by the **talent dispersion** λ:

- the range of talent is large enough,

$$\overline{x}_F > c - \pi - \mu_q(0.5) \qquad \text{and} \qquad \underline{x}_F < c - \pi - \bar{M}$$

  for $\lambda' \geq \lambda$ and $q' \geq q$.
- the support of F_λ weakly increases in the set-inclusion sense, with $\lim_{\lambda \to \infty} \bar{x}_{F_\lambda} = \infty$.
- for any x that is (eventually) inside the support of F_λ, we assume that the hazard rate converges to zero pointwise,

$$\lim_{\lambda \to \infty} \frac{F'_\lambda(x)}{1 - F_\lambda(x)} = 0.$$

Steady States (1/3)

Proposition

(a) *The economy admits two homogeneous steady states* $\hat{\phi} \in \{0, 1\}$ *if and only if the most able individuals require some mentorship boost to participate,*

$$c - \bar{x}_F - \mu(0) \geq \pi.$$

The homogeneous steady states are stable whenever the inequality is strict.

(b) *The economy always admits a mixed steady state* $\hat{\phi} \in (0, 1)$.

- There can even be steady states where the workforce is dominated by the population minority, $\hat{\phi} < 0.5$.
- Example: In South Africa,80% of the economically active population is Black African, yet they still hold only 14.3% of top management jobs even over 20 years after the end of apartheid (BBC News, 2019).
- Our analysis suggests that mentorship disparities can sustain such a bias towards the minority indefinitely.

Steady States (2/3)

Proposition

(c) *If $b > 0.5$, any majority-dominant workforce converges to a steady state that is biased towards the majority.*

- First, we show that a majority-dominant senior workforce always attracts more juniors of the majority than of the minority.
- Because talent is equally distributed, a fair steady-state labor market with $\hat{x}_1(\phi) = \hat{x}_2(\phi)$ would therefore mirror the composition of the population, $\hat{\phi} = b$.
- We show that the majority will eventually be over-represented ($\hat{\phi} > b$) by ruling out any steady states with composition $\hat{\phi} \in (0.5, b]$.

Steady States (3/3)

Proposition

(d) *As mentor capacity $q \to \infty$, the economy admits a stable steady state near b. The steady states of the economy tend towards $\{0, b, 1\}$.*

(d) *As talent dispersion $\lambda \to \infty$, the economy admits a stable steady state near b. The steady states of the economy tend towards $\{b\}$.*

- If either mentor capacity or talent concentration is large, labor supply hardly responds to differences in mentor availability:
  - If q is large, this is because even a small representation yields a near-maximal mentorship boost;
  - if λ is large, this is because there are very few juniors in the middle of the talent distribution who could be swayed to participate with mentorship.

  Thus, for any interior ϕ, the ratio of group-1 to group-2 individuals with talent above $\hat{x}_i(\phi)$ converges to $b : 1 - b$, ruling out any other steady states.

Welfare-maximizing steady state

Proposition (Optimal Steady State)

For sufficiently large mentor capacity q or high talent dispersion λ, the surplus-maximizing (stable) steady state is nearly fair, $\phi^*_{SS} \approx b$.

- As mentor capacity increases, even a handful of minority mentors can provide a near-perfect boost to minority juniors.
- As a result, the efficiency tension resolves in favor of talent recruitment, and surplus is maximized at a nearly fair steady state.
- Highly concentrated talent makes all other workforce compositions unsteady.
- Although the mixed steady states tend towards fairness, the minority remains *under*represented at the mixed steady state in the sense that $\hat{\phi} > b$ for any finite q or λ.

Optimal long-run intervention (1/5)

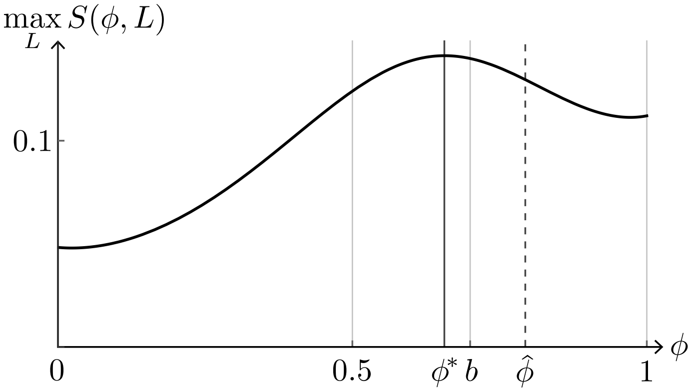


Figure 1: Social surplus as a function of labor force composition ϕ.

Optimal long-run intervention (2/5)

Proposition (Optimal Intervention)

The optimal labor force composition ϕ^* depends on mentor capacity q as follows:

(a) *If $c - \pi > \bar{x}_F$ and q is sufficiently small, a homogeneous labor force is optimal.*

(b) *If $b > 0.5$, the optimal labor force is always dominated by the majority, $\phi^* > 0.5$. However, as long as mentor capacity is sufficiently large, $q > Q$, the optimal labor force is biased in favor of the minority, $\phi^* \in (0.5, b)$.*

- Including even just the most talented minority members dilutes mentoring for the majority, and this effect can outweigh for small mentor capacities.
- However, that a homogeneous labor force is never optimal when the upper bound on talent, $\bar{x}_F$, is large enough.
- Larger mentor capacity q makes the mentoring dilution less costly for the majority.
  $\rightarrow$ optimal labor market is biased in favor of the minority

Sketch of proof for over-representation of minority

- $\phi^* > 0.5$ is straightforward.
- For sufficiently large q, we can show that if $\phi > b$, $S^*(b) > S^*(\phi)$ where $S^*(\phi) = \max_{L \geq 0} S(\phi, L)$.
- The reason is that with large mentor capacity, using the talent of the minority outweights the mentoring benefit and
- The marginal mentoring benefit of an additional minority mentor is higher than of a marginal majority mentor for sufficiently large q.

Optimal long-run intervention (3/5)

Proposition (Optimal Intervention)

(c) *The optimal composition converges to that of the population* $\lim_{q \to \infty} \phi^* = b$.

- The size of the bias disappears in the limit $q \to \infty$.
- In such a market, the policy maker recruits minority workers with talent **below** the marginal majority worker – not just as a transitory course correction, but as an ongoing policy.
- **Intuition:** Workers don't internalize their own positive mentoring externality on future generations. When mentors are efficient (q large), the social returns warrant minority subsidies that exceed the inherent mentoring advantage of the majority.

Optimal long-run intervention (4/5)

Proposition (Optimal Intervention)

(c) *For large enough talent dispersion λ, the surplus-maximizing economy is biased towards the minority (majority) whenever $M_q'(b) < 0$ ($M_q'(b) > 0$).*

- If talent is sufficiently important, the shape of the total surplus generated by mentorship $M_q(\phi) := \phi\mu_q(\phi) + (1-\phi)\mu_q(1-\phi)$ determines whether the majority or the minority is over-represented in the economy at the optimal constant labor force.
- In particular, if $\mu(b) + b\mu'(b) < \mu(1-b) + (1-b)\mu'(1-b)$, then overrepresentation of the minority increases the gains from mentoring.
- Recall, that we assumed that $M_q'(\phi) < 0$ for all $q \geq Q_\phi$.

Optimal long-run intervention (5/5)

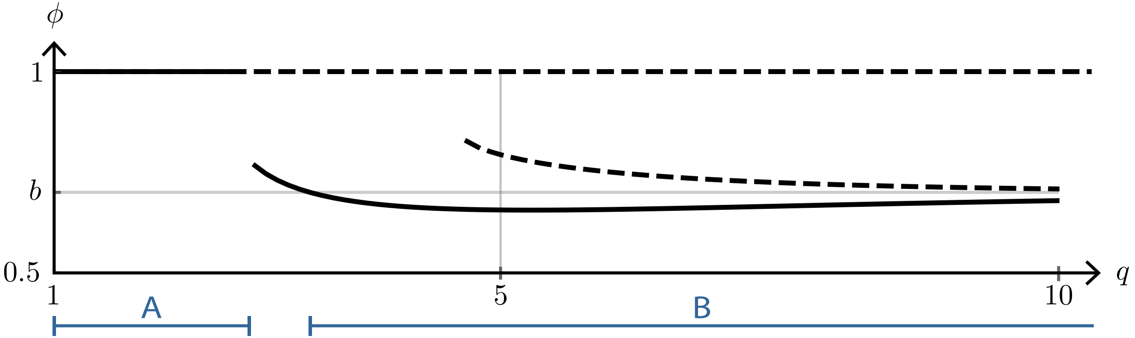


Figure 2: Optimal composition as a function of mentor capacity q (solid line). Stable steady state compositions are indicated with dashed lines.

Policy Instruments: Educational Incentives

- Let $\boldsymbol{\Delta} \in \mathbb{R}^2$ denote a transfer schedule where Δ_i represents the net transfer to individuals in group i. (available to *all* interested minority students)
- Because the labor market remains unrestricted, so $w = \pi$.
- With the intervention:

$$\Delta_i = c - \pi - \tilde{\mu}(L_i, L_{\neg i}, L_i^*, L_{\neg i}^*) - x_i^* \qquad \forall i = 1, 2,$$

  the status quo labor force becomes $\boldsymbol{L}^*$ after one period, but the policy needs to stay in effect since $\boldsymbol{L}^*$ is generally not a steady state.
- Once $\boldsymbol{L}^*$ is reached, it can be maintained in a way that is budget-balanced:

$$0 = \phi^* \Delta_1 + (1 - \phi^*)\Delta_2.$$

Ability-based fellowships *only* affect the extensive margin if the available pool exceeds the unregulated student supply!

Policy Instruments: Labor Force Quotas - free wages

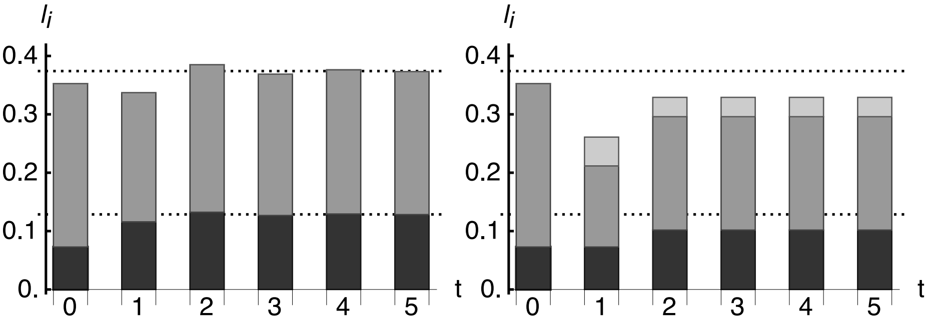
- Oversupply of educated group-i workres would drive wage w_i to zero.
- The size of the cohort ℓ and the market wages w_i are then uniquely determined by the market clearing equations

$$\left\{ \begin{aligned} \phi^*\ell =& \quad\quad b\,(1-F(c-\tilde{\mu}(L_1,L_2,\phi^*\ell,(1-\phi^*)\ell)-w_1)) \\ (1-\phi^*)\ell =& \, (1-b)\,(1-F(c-\tilde{\mu}(L_2,L_1,(1-\phi^*)\ell,\phi^*\ell)-w_2)) \\ \pi =& \, \phi^* w_1 + (1-\phi^*)w_2. \end{aligned} \right.$$

- Same as budget balance condition where the subsidy is $\boldsymbol{\Delta} = \boldsymbol{\pi} - \boldsymbol{w}$.
- This implies that a binding quota *raises* minority and *depresses* majority earnings relative to the unconstrained market, $w_1 < \pi < w_2$.

Policy Instruments: Labor Force Quotas - fixed wages

- Let us impose $w_1 = w_2$.
- Then, $w_1 = w_2 = \pi$.
- Majority worker demand is then capped at $\ell_1 = \frac{\phi^*}{1-\phi^*}\ell_2$, while all minority workers are hired. $\rightarrow$ Job insecurity.



Policy Instruments: Mentor Training

- For large mentor capacity q, the tension between talent recruitment and mentorship efficiency disappears.
- Thus, the need for long-run market intervention disappears if mentorship itself can be improved through cross-group exposure, mentor training, and networking support for minority youth.

Summary

- Long-term workforce composition when mentoring effectiveness is group specific and innate talent is important.
- Source of inefficiency: Juniors do not internalize their positive mentoring externality when making education decisions.
- Naturally, a workforce with $> 50\%$ converges to a steady state where the majority is over-represented.
- Long-run interventions can improve welfare:
  - If the mentor capacity is sufficiently large, then the minority should be even over-represented.
  - Implementation through group-specific fellowships to include minority workers of lower innate talent.
- A fair labor market composition is close to efficient if the mentor capacity is sufficiently large.

Thank you!