Affirmative Action in Higher Education
Allowing for Strategic School Quality and Student Effort Choice

Saattvic[1] [2]

March, 2025

CLICK HERE FOR LATEST VERSION

Abstract:

Affirmative action policies in higher education aim to reduce economic inequality between historically privileged and underprivileged groups. This paper theoretically assess the effect of affirmative action policies on schools' quality choice and students' effort choice. Students from privileged and underprivileged groups are endowed with ability and parental human capital, which along with school quality and effort produce human capital measured by grades. Students who meet a grade threshold attend university and further increase their human capital. In the baseline model school quality and student effort are held fixed. Without affirmative action, privileged group students are overrepresented at university and achieve a higher average realized human capital. Affirmative action reduces but does not eliminate inequality between groups. When schools can choose quality, early movers compete for privileged group students by investing highly in quality, while late movers invest less in quality and educate underprivileged group students. Affirmative action makes underprivileged group students more attractive and weakens competition for privileged group students, resulting in a lower quality chosen by early movers. The effect on late movers' quality choice is ambiguous, as they adjust quality to put more density in the vicinity of the threshold grade in an effort to take advantage of the affirmative action policy. When students choose effort levels, those with low endowment choose low effort levels and do not attend university, those with high endowment choose medium effort levels and attend university comfortably exceeding the threshold grade, and those with medium endowment choose high effort levels to meet the threshold grade and attend university. Affirmative action reduces effort by those privileged group students who lose university seats, but increases effort for others who must now meet a higher threshold. The opposite is true for underprivileged students, and the net effects on average effort are ambiguous for both groups. Given these ambiguities in school quality and student effort choice, it is possible for affirmative action to cause inequality to fall, remain unchanged, or increase.

[1] saattvic@sauder.ubc.ca, Sauder School of Business, University of British Columbia

[2] I carried out initial work on this paper under the supervision of Margaret Stevens while I read for an MPhil in Economics at the University of Oxford. I thank Margaret Stevens and two unknown reviewers for useful comments and insights. I acknowledge funding from the Bharat Petroleum Scholarship while I was at the University of Oxford. I acknowldge funding from the following scholarships and grants received during my time at the University of British Columbia: Canada Graduate Scholarships – Doctoral, MINDS Initiative Doctoral Awards, Joint Initiative for Digital Citizen Research Doctoral Scholarship, UBC Four Year Fellowships, UBC Sauder Dean Earle McPhee Fellowship in Commerce and Business Administration, UBC Sauder President's Academic Excellence initiative PhD Awards, and Sauder School of Business Graduate Awards.

1 Introduction

Affirmative action policies in higher education are often used as a tool to uplift economically lagging groups which have historically faced discrimination. India and South Africa have explicit quotas, while universities in the US and other developed countries have long stated a preference for higher diversity and inclusion of historically marginalized groups. Such policies can be controversial, as the protests against admissions quotas in India in 1990, or the US administration's steps to roll back diversity, equity and inclusion initiatives in 2025 show. Proponents invoke social justice and other arguments involving externalities that affirmative action policies can help solve; while opponents argue that such policies encourage slack and are ultimately ineffective.

This paper theoretically investigates the effect of affirmative action policies on schools' quality and students' effort level choices, and the consequent implications for the effectiveness of affirmative action policies in reducing inequality between groups.

My model has two groups – the (historically) privileged and (historically) underprivileged. Students from each group are endowed with intrinsic ability, which is distributed similarly across groups, and parental human capital, where the distribution for the privileged group dominates that of the underprivileged group. Both variables, along with school quality and effort level, are inputs into a human capital production function, which results in a grade. University admissions follow a simple threshold rule – students with grades above a threshold attend university and further increase their human capital, with the threshold being determined so as to exactly clear college seats. Students maximize realized human capital in their decision making, while schools maximize the realized human capital of their students less the costs of investing in quality. Affirmative action policies are designed to achieve equal rates of representation at university across groups

To derive baseline results, I hold school quality and student effort levels fixed and homogeneous across all students. Since parental human capital is an input into grades, the privileged group's grade distribution dominates that of the underprivileged group, resulting in overrepresentation at university. A higher average parental capital and overrepresentation at university both push average realized human capital for the privileged group higher than for the underprivileged group. Affirmative action achieves equal rates of representation at university by increasing the threshold grade for privileged group students and decreasing it for underprivileged group students. The difference in average realized human capital across groups falls but remains positive because of the disparity in parental human capital.

I then let schools choose quality sequentially. Without affirmative action, schools prefer privileged to underprivileged group students because they are likelier to have higher grades and therefore more likely to go to university. All students prefer schools with higher quality. Thus, early movers invest in high quality in a bid to attract privileged group students, up to the point where the payoff is the same as that of investing in lower quality and settling for an underprivileged group student. Late movers, having lost the opportunity to attract privileged group students, invest in lower quality and educate underprivileged group students. Compared to the baseline, the disparity between grade distributions, the extent of overrepresentation at university, and the difference between average realized human capital across groups are all exacerbated by privileged students attending better quality schools. Affirmative action increases the probability of underprivileged students attending university, which makes them more attractive to schools. Early mover schools therefore compete less intensely to attract privileged group students and invest less in quality. However, the effect on the quality choice of late movers is ambiguous – schools adjust quality so as to put more probability mass in the grade distribution in the vicinity of the threshold grade, so as to increase the probability of their students attending university as a result of affirmative action. Since late movers' quality could also reduce, inequality in average realized human capital across groups could increase.

Finally, I let students choose effort levels (while reverting to constant and homogeneous school quality choice). The competition for university seats leads to optimal effort choices that are discontinuous in endowment levels. For low endowment levels, students find that working extra hard to meet the threshold grade is costly enough to offset the increase in human capital from attending university, and so choose not to attend university and to exert low levels of effort. For high endowment levels, students find that moderate effort levels comfortably put them above the threshold grade. For moderate endowment levels, students find that the optimal effort level conditional on attending university would place them below the threshold, so they choose high effort levels to meet the threshold. Unlike for those with low endowments, these students find that the increase in human capital from attending university outweighs the costs of extra effort. Affirmative action reduces the threshold grade for the underprivileged group and increases it for the privileged group. Within the underprivileged group, some who earlier chose to not attend university now find it profitable to exert more effort to meet the lower threshold grade. However, others who earlier worked hard to meet the threshold grade now need to exert lower effort. Similarly, for the privileged group, some students no longer find it worthwhile to work hard to meet the increased threshold and reduce their effort levels, while others must raise effort levels to attend university. The effect on average effort is ambiguous for both groups, leading to the possibility of increasing inequality in average realized human capital.

This paper adds to the theoretical literature on affirmative action in higher education[1] by incorporating strategic decision making by schools and students. It also contributes possible explanations for the empirical literature that finds affirmative action policies have little to no effect on eventual earnings and inequality.[2]

2 Baseline Model

2.1 Model Set Up

Let there be a population of n students, a fraction λ of whom come from a historically privileged community (denoted P) and a fraction $(1-\lambda)$ come from a historically underprivileged community (denoted U).

Students are endowed with ability $a \in [0, \bar{a}]$ and parental capital $p \in [0, \bar{p}]$. For $i \in \{P, U\}$, let $H_i(a, p)$

[1] Durlauf (2008); De Fraja (2002); Rotthoff (2008); del Rey Canteli and Racionero (2008); Epple et al. (2008); Fryer Jr et al. (2008).

[2] Bertrand et al. (2010); Arcidiacono (2005).

be the joint distribution, $X_i(a)$ the marginal distribution of ability, and $B_i(p|a)$ the conditional distribution of parental capital given ability.

Assumption 2.1. *(i) $X_P(a) = X_U(a)$ $\forall a$; (ii) $B_P(p|a) < B_U(p|a)$ $\forall a, p$; and (iii) $X_i(a)$ and $B_i(p|a)$ are continuous.*

Thus, ability is distributed the same across groups, but P students are likely to are likely to have higher parental capital for a given ability level. The law of total probability implies that the marginal distribution of parental capital for the P group (strictly) first order stochastically dominates that of the U group. This assumption reflects persisting inequalities in capital between historically privileged and disadvantaged groups.

Students study in schools, and earn grades according to the production function

$$g = g(a, p, q, e), \tag{1}$$

where $q \geq 0$ is school quality and $e \geq 0$ is effort.

Assumption 2.2. *(i) $g(\cdot,\cdot,\cdot,\cdot)$ is continuous and twice continuously differentiable; (ii) g is increasing in all its arguments; (iii) all second order cross derivatives are positive; and (iv) $g(0,\cdot,\cdot,\cdot) = g(\cdot,0,\cdot,\cdot) = g(\cdot,\cdot,0,\cdot) = g(\cdot,\cdot,\cdot,0) = 0$.*

(i) is a simplifying assumption, (ii) implies higher inputs lead to higher outputs, (iii) implies inputs are complementary, which is a common modeling assumption, and (iv) fixes levels and ensures grades are non-negative for convenience. p entering the grades function in this way reflects traditional mechanisms for the inter-generational transfer of human capital. Parents with more capital can improve their children's school outcomes by, for example, providing better home teaching, hiring better home tutors, providing better infrastructure, etc.

Based on their grades, students compete for $\bar{A} < n$ university seats. College admissions follow a simple threshold policy – the $\bar{A}$ students with the highest grades are admitted.

The payoff for students is their human capital at the end of this game. For those not admitted to university, human capital is g. Those admitted to college accumulate human capital according to the production function $k(g)$.

Assumption 2.3. *(i) $k(\cdot)$ is continuous and twice differentiable; and (ii) $\frac{\partial k(g)}{\partial g} \geq 1$ and $k(g) \geq g$ $\forall g$.*

(ii) implies that a university education never reduces human capital, and students with higher grades benefit at least as much as those with lower grades.[3] However, this does not mean that students with higher grades are 'better able' to take advantage of a university education since it is consistent with university increasing human capital by a fixed proportion.[4] The assumption is even consistent with a constant increase in human capital.

[3] Bertrand et al. (2010) reports that going to engineering college increased incomes of low-caste college seat gainers less than for displaced high-caste students.

[4] Data in Bertrand et al. (2010) support a proportionate increase of 40-70%.

Since the primary motivation of the paper is to investigate the effects of strategic school quality and student effort choice, to derive baseline results I assume these are both non-strategic.

Assumption 2.4. *(i) q is non-stochastic and constant across all schools; and (ii) e is non-stochastic and constant across all students.*

2.2 Outcome without Affirmative Action

Since there is no strategic decision making by any agents in the baseline model, I do not need to analyze any optimization decision. The outcome flows mechanically.

Lemma 2.1. *The baseline outcome without affirmative action has the following characteristics.*

(a) *Grades for group i are distributed according to $F_i(g) = \mathbb{E}_a [B_i(P(a|g))]$ where $P(a|g)$ is implicitly defined by $g = g(a, P(a|g), q, e)$.*

(b) $F_P(g) < F_U(g)\ \forall g$.

(c) *The threshold grade g^* for university admissions is given implicitly by $[(1-F_P(g^*))\lambda+(1-F_U(g^*))(1-\lambda)]n = \bar{A}$.*

(d) *The P group is overrepresented in university.*

(e) *Average realized human capital is higher for the P group.*

Proof. See Appendix. □

Since grades are influenced by parental capital, and since the P group's parental capital distribution dominates that of the U group, grades for P students dominate those of U students as well. Under a threshold admissions system, this translates into P students being overrepresented in university, in that the proportion of university students made up of P students exceeds their share of the general population. Since a higher proportion of P students go to university and since grades influence eventual human capital attainment, P students' distribution of eventual human capital dominates that of U students', resulting in a disparity in average human capital attainment across groups.

2.3 Imposing Affirmative Action

Definition 2.1 (Affirmative Action)**.** *Affirmative action is a policy that achieves equal representation, i.e. the proportions of P and U students admitted to university must be the same.*

Lemma 2.2. *Compared to the outcome without affirmative action,*

(a) *Threshold grades rise for P students and fall for U students.*[5]

(b) *Average realized human capital rises for U students and falls for P students.*

(c) *The difference in average realized human capital for P and U students remains positive.*

[5] Bertrand et al. (2010) report threshold grades for engineering institutions covered in the study were 480/800 for upper caste students and 182/800 for lower caste students.

(d) Economy-wide average realized human capital weakly falls.[6]

Proof. See Appendix. □

Affirmative action reduces P representation and increases U representation at university. This can only be achieved by reducing the threshold grade for U students so that more are let in, and increasing it for P students so that fewer are let in. This directly reduces average realized human capital for the P group since a subset goes from attending to not attending university, and vice versa for the U group. However, affirmative action does not eliminate inequality because even though similar proportions of P and U groups go to university, the grades distribution for the P group still dominates that for the U group, the P students who attend university have on average higher grades than their U group colleagues, and students with higher grades gain at least as much from university as those with lower grades. Economy-wide average realized human capital either stays the same or falls because the U students who gain access to university have lower grades than the P students who lose access, and students with higher grades gain at least as much from university as those with lower grades.

3 School Quality Choice

Proponents often argue that affirmative action is necessary to counterbalance differences in the quality of schooling available to historically privileged and underprivileged groups. In this section, I allow schools to choose quality and loosely recreate a stylized fact – privileged group students tend to attend higher quality schools. I then analyze the effect of affirmative action policies on school quality choice, and the consequent implications for inequality in realized human capital.

3.1 Model Set Up

Assumptions 2.1, 2.2, 2.3 and 2.4(ii) continue to hold. However, schools are now decision making agents. Let there be n schools, each with the capacity to educate one student.[7] Students (or their parents on their behalf) also become decision making agents. The sequence of actions is as follows.

1. n students are born, with nature assigning each person into group P with probability λ and group U with probability $1 - \lambda$.
2. Schools (indexed by j) choose publicly observable quality q_j sequentially in a random sequence.[8]
3. Schools observe only each student's group affiliation (not a or p) and make an offer to a student. If the offer is rejected, the school makes another offer until the seat is filled.
4. Students choose which offer to accept.

[6] Bertrand et al. (2010) conclude that affirmative action policies come at an absolute cost.

[7] The analysis would be materially unchanged but more cumbersome with a smaller number of schools each with the capacity to educate several students.

[8] A version of this game with simultaneous entry and quality choice does not have a pure strategy equilibrium, but the more cumbersome mixed strategy equilibrium achieves materially the same outcome.

5. Students undergo schooling and possibly university, and receive payoffs (realized human capital).

Assumption 3.1. *All schools are a priori symmetric with the following characteristics*

(i) *Costs of increasing quality are* $C(q)$ *with* $C(0) = 0$, $C'(q) > 0$ *and* $C''(q) > 0$

(ii) *The payoff function is* $S(q|i, g^*(i)) = \mathbb{E}\left[\text{pupil's realized human capital } |i, g^*(i)\right] - C(q)$, *where* $g^*(i)$ *returns the threshold grade for group* i

(iii) $\{q : S(q|i, g^*(i)) \geq 0\} \neq \phi$

(iv) $\exists\ \bar{q}_i$ *such that* $S(q|i, g^*(i)) < 0\ \forall q \geq \bar{q}_i$ *for* $i \in \{P, u\}$

(ii) postulates that schools look to the expected realized human capital of their pupils. This is motivated by schools looking to enhance their reputation, which can result in higher fees or donations from alumni. (iii) and (iv) together ensure that schools will want to operate.

3.2 Outcome without Affirmative Action

Lemma 3.1. *In equilibrium*

(a) *Grades for group* i *are distributed according to* $F_i(g|q) = \mathbb{E}_a\left[B_i(P(a|g, q))\right]$ *where* $P(a; g, q)$ *is implicitly defined by* $g = g(a, P(a; g, q), q, e)$.

(b) $F_P(g|q) < F_U(g|q)\ \forall g$.

(c) $F_i(g|q'') \leq F_i(g|q')$ *when* $q'' \geq q'$.

(d) *The first* λn *schools choose 'good' quality* $q_G = \max\{q : S(q|P, g^*) = \max_q S(q|U, g^*)\}$.

(e) *The remaining schools choose 'bad' quality* $q_B \in \arg\max_q S(q|U, g^*) < q_G$.

(f) P *students attend good quality schools and* U *students attend bad quality schools.*

(g) P *students are overrepresented at university.*

(h) *Average realized human capital is higher for the* P *group.*

Proof. See Appendix. □

Analogous to the baseline model, for a given school quality level, the grades distribution for P students (strictly) dominates that of U students because parental capital influences grades. Moreover, since school quality influences grades, the grades distribution associated with a higher quality school dominates that associated with a lower quality school for a student from a given group.

The Appendix shows that the representative school's payoff can be written as

$$S(q|i, g^*) = \mathbb{E}_{F_i|q}\left[g\right] + \int_{g^*}^{\bar{g}} (k(g) - g) dF_i(g|q) - C(q) \tag{2}$$

$S(q|i, g^*)$ is continuous since all its constituent functions and probability distributions are continuous. Moreover, $S(0|i, g^*) = 0$ since $g(a, p, 0, e) = 0$; there is a set of values for which it is positive, and it eventually turns negative. Also, $S(q|P, g^*) > S(q|U, g^*)$, i.e. for the same quality chosen, educating a P student will give a higher payoff than educating a U student. This follows from P students' grade distribution dominating that of U students, and P students being more likely to go to university. Figure 1 depicts a configuration of $S(q|P, g^*)$ and $S(q|U, g^*)$ curves that obey these conditions. q_B is defined as the quality that maximizes the payoff from educating a U student for sure, and q_G as the highest quality that gives the same payoff from educating a P student as q_B gives from educating a U student.

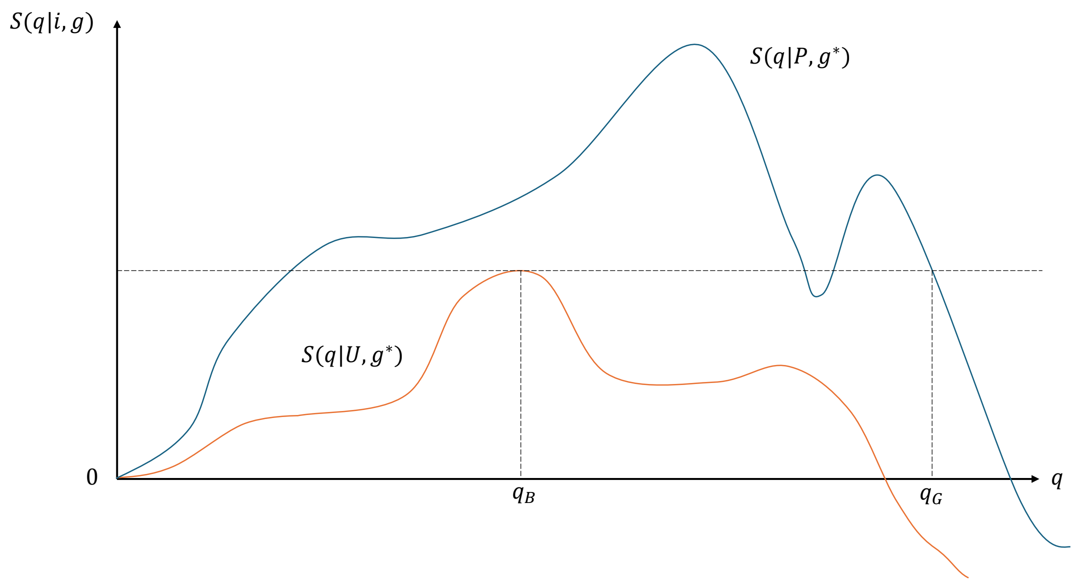

Figure 1: $S(q|P, g^*)$ and $S(q|U, g^*)$

Consider the best response strategies for schools and students in the game defined earlier. Students prefer to attend better quality schools because it increases their realized human capital. Moreover, schools associate P students with a higher payoff for any quality choice. The Appendix shows that the best response strategies for schools are as follows.

- For the first λn schools, choose q_G
- For school k within the $(\lambda n+1)$-th to $(n-1)$-th schools, if at least λn schools have chosen at least q_G, choose q_B; if fewer than λn schools have chosen at least q_G, if the minimum of the top $\lambda n - k$ qualities chosen so far (defined as q') is less than q_G choose $\arg\max_{q\in(q',q_G)} S(q|P, g^*)$, else choose q_G
- for the last school, if the minimum of of the top $\lambda n - 1$ qualities so far (defined as q') is below q_G, choose $\arg\max_{q\in(q',q_G)} S(q|P, g^*)$, else choose q_B

The outcomes described in Lemma 3.1(d)-(f) are part of a subgame perfect equilibrium because there are no profitable deviations. For any of the first λn schools choosing $q > q_G$ will attract a P student for sure but lead to a lower payoff; and choosing $q < q_G$ will attract a U student for sure (since a later school could replace it in the top λn) and lead to a weakly lower payoff. For any of the remaining schools, attracting a P student would require choosing $q \geq q_G$, which leads to a weakly lower payoff than attracting a U student with quality q_B; and choosing any $q < q_G$ other than q_B will attract a U student for sure and lead to a lower payoff. Students prefer good quality schools over bad quality schools, and since each school makes offers first to P students, they are distributed among good quality schools, leaving bad quality schools to educate U students.

P students' overrepresentation at university is even starker in this model compared to the baseline. Their grade distribution dominates that of U students not only because of the difference in parental capital, but because they receive a better education at school. This overrepresentation contributes to a higher average realized human capital for the P group.

3.3 Imposing Affirmative Action

Assumption 3.2. $g^*_P > g^* > g^*_U$

This assumption rules out counterintuitive equilibria that are not already ruled out given the limited functional form and distributional assumptions so far.

Lemma 3.2. *Compared to the outcome without affirmative action,*

(a) The first λn schools choose 'good' quality $q'_G = \max\{q : S(q|P, g^*_P) = \max_q S(q|U, g^*_U)\}$

(b) The remaining schools choose 'bad' quality $q'_B \in \arg\max_q S(q|U, g^*_U) < q'_G$

(c) P students attend good quality schools and U students attend bad quality schools.

(d) Good quality schools lower their quality but the effect on the quality of bad quality schools is ambiguous.

(e) The effect on the difference in average realized human capital is ambiguous, but it remains positive.

(f) The effect on economy-wide average realized human capital is ambiguous.

Proof. See Appendix. □

Since the threshold grade falls for U students and rises for P students, at a given q, the expected payoff for schools rises from educating a U student and falls from educating a P student. Consequently, $S(q|U, g^*_U) > S(q|U, g^*)$ and $S(q|P, g^*_P) < S(q|P, g^*)$. However, the Appendix shows that $S(q|P, g^*_P) > S(q|U, g^*_U)$, i.e. schools still prefer P students to U students at every q. This is because even though the probability of a student going to university is now equal across groups, P students' grade distribution still dominates that for U students for a given q and therefore P students will likely have a higher level of realized human capital. Figure 2 shows a possible configuration of how payoffs for schools change.

The best response strategies and the resulting equilibrium do not qualitatively change. Students still prefer better quality schools to worse ones, and schools still prefer P students to U students. Since the curve for P students is still above that for U students, the first λn schools still find it best to choose the highest quality that will attract a P student that also gives at least as much payoff the best possible payoff from educating a U student, and the remaining schools choose the quality that maximizes the payoff from educating a U student. P Students still attend high quality schools and U students still attend low quality schools.

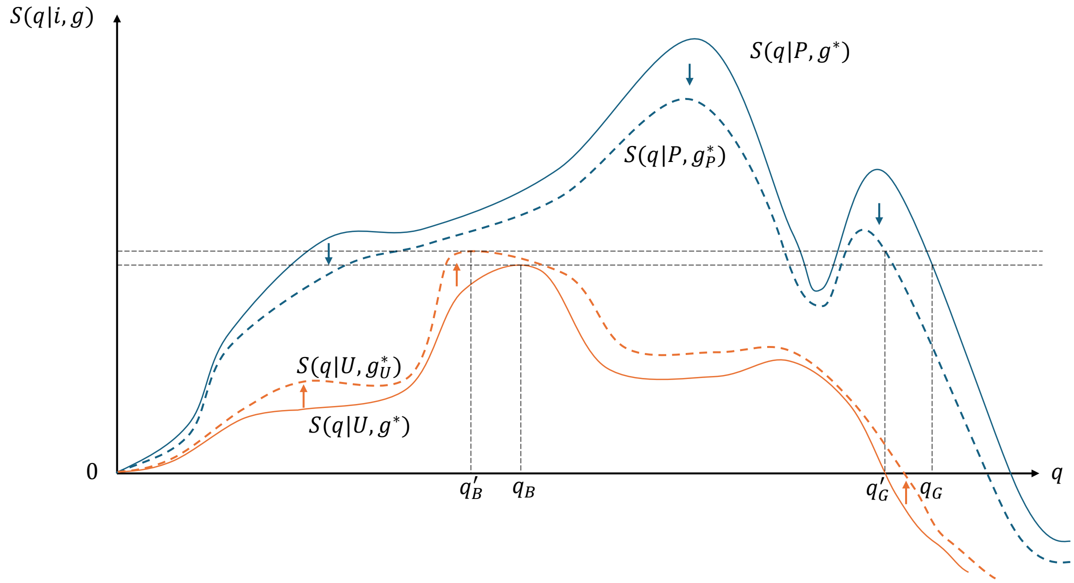

Figure 2: Affirmative Action's Effect on School Quality Choice

$q'_G < q_G$ unambiguously for two reasons. First, since the probability of P students going to university falls, even if the attractiveness of educating a U as an alternative does not change, there is less incentive to invest in quality. Second, the alternative of educating a U student becomes more attractive, so schools are even less willing to invest to attract a P student.

The change in the quality of low quality schools is ambiguous. q'_B is the quality level that maximizes $S(q|U, g^*_U)$, so whether it is more or less than q_B depends on how the objective function shifts as the threshold grade changes. The Appendix applies a monotone comparative static result from Milgrom and Shannon (1994) and shows that

$$\frac{\partial f_U(g|q)}{\partial q} \leq 0 \Rightarrow q'_B \leq q_B \qquad \& \qquad \frac{\partial f_U(g|q)}{\partial q} \geq 0 \Rightarrow q'_B \geq q_B, \tag{3}$$

or that if $f_U(g|q)$ is decreasing in q in the vicinity of g^* then quality falls and vice versa.

Figure 3 helps understand the intuition behind this result. Consider a single peaked density function $f_U(g|q)$ at some quality q'. If quality increased to some q'', then by Lemma 3.1(c) the new distribution will stochastically dominate the old one. In the diagram, the new density is depicted by a dashed line. If the threshold grade is at g_1, an increase in quality will increase the probability mass in the vicinity of the threshold grade. In this scenario, a reduction in threshold grade would increase the probability of a student getting into university much faster. However, if the threshold is at a point (like g_0) where an increase in quality reduces probability mass, then increasing quality reduces the effectiveness of a falling threshold in

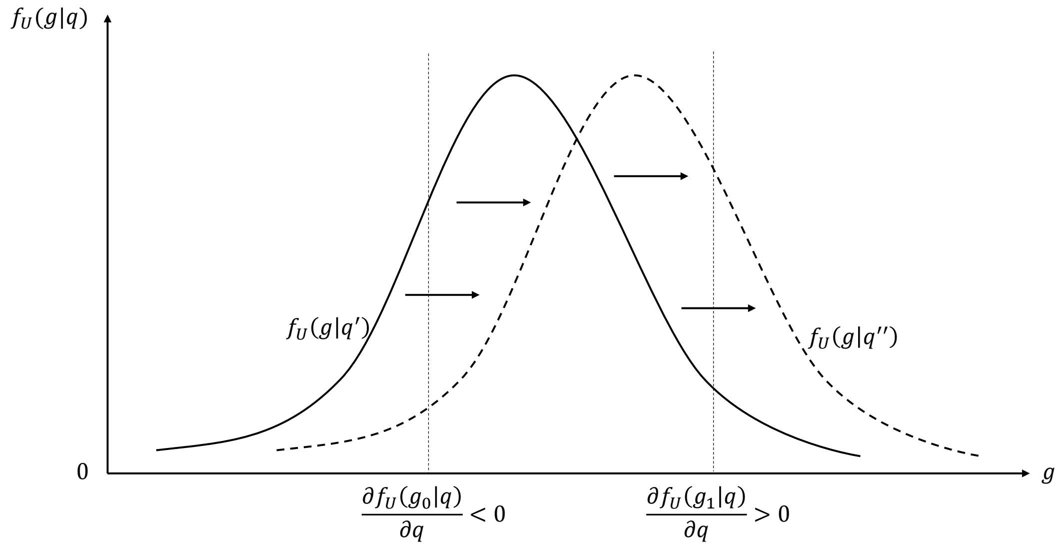

Figure 3: The Effect of $f_U(g|q)$ on School Quality Choice

increasing the probability of going to university. The incentive at this grade is to *reduce* quality to put more probability mass in the vicinity of the threshold, so as to take maximum advantage of the externally imposed reduction in threshold. Thus, schools adjust quality to put more people in the vicinity of the threshold, so as to better exploit the effect a falling threshold has on increasing the prospects that the student goes to university.

Average realized human capital for the P group remains higher than for the U group because even though they attend university at similar rates, the distribution of P student's parental human capital still dominates that of U students, and P students continue to attend better quality schools. Since bad quality schools might lower quality, the effect of affirmative action on the average realized human capital for the U group is ambiguous even though their propensity of attending university increases. This leads to the following proposition.

Proposition 3.1. *Affirmative action may cause schools educating U students to reduce quality, which might in turn reduce their average realized human capital and might increase the difference between average realized human capital across groups.*

4 Student Effort Choice

Opponents of affirmative action often criticize members of the beneficiary group as slackers who take advantage of society's benevolence. In this section I abstract from school quality choice but allow students to choose effort levels. I find that competition for college seats leads to some students from both groups working hard. I then analyze how affirmative action changes the relative numbers of hard workers in both groups, and the consequent implications for inequality in realized human capital.

4.1 Model Set Up

Assumptions 2.1, 2.2, 2.3 and 2.4(i) continue to hold. Students are now decision making agents.

Assumption 4.1. *The following apply.*

(i) *a and p enter $g(a,p,q,e)$ in a way such that $g(a,p,q,e) = g(\tau(a,p),q,e)$ where $\tau(a,p)$ is increasing in both arguments.*

(ii) *Students maximize their realized human capital by choosing effort.*

(iii) *Effort is costly, with costs of effort being continuous, twice differentiable and homogeneous across students at $C(e)$, with $C(0) = 0$, $C'(e) > 0$ and $C''(e) > \max\left\{\frac{\partial^2 g}{\partial e^2}, \frac{\partial k}{\partial g}\frac{\partial^2 g}{\partial e^2} + \left(\frac{\partial g}{\partial e}\right)^2 \frac{\partial^2 k}{\partial g^2}\right\}$.*

(iv) $\frac{\partial^2 k}{\partial g^2} \geq -\left(\frac{\partial g}{\partial e}\frac{\partial g}{\partial \tau}\right)^{-1} \frac{\partial k}{\partial g}\frac{\partial^2 g}{\partial \tau \partial e}$.

(i) is a separability assumption on $g(a,p,q,e)$ that allows the condensing of variation in individual characteristics into one 'educability' variable. For instance, if $g(a,p,q,e) = apqe$, then $\tau(a,p) = ap$ and $g(\tau,q,e) = \tau qe$. This assumption is made for convenience to facilitate two-dimensional graphical analysis. The last condition in (iii) ensures an interior solution for effort. (iv) specifies that $k(g)$ is not too concave, which significantly simplifies the analysis below.

The representative student's payoff is thus defined by

$$S(e|\tau) = \begin{cases} g(\tau,q,e) - C(e) & \text{if student does not attend university} \\ k(g(\tau,q,e)) - C(e) & \text{if student attends university} \end{cases} \tag{4}$$

Since students know τ and q is a known constant, students can calculate their realized human capital precisely.

Lemma 4.1. *The following hold in relation to τ.*

(a) *$g(\tau,q,e)$ is increasing in τ and all cross derivatives are positive.*

(b) *τ for group i is distributed according to $F_i(\tau) = \mathbb{E}_a\left[B_i(P(a|\tau))\right]$ where $P(a|\tau)$ is implicitly defined by $\tau = \tau(a, P(\tau|a))$.*

(c) *$F_P(\tau) < F_U(\tau)\ \forall\tau$.*

Proof. See Appendix. □

4.2 Outcome without Affirmative Action

Lemma 4.2. *In equilibrium*

(a) The optimal effort schedule $e(\tau)$ satisfies

$$\begin{aligned} &\frac{\partial g(\tau,q,e(\tau))}{\partial e} = C'(e(\tau)) && \textit{when } \tau < \tau^* \\ &g(\tau,q,e(\tau)) = g(\tau^*,q,e(\tau^*)) \equiv g^* && \textit{when } \tau^* \leq \tau \leq \tau_0 \\ &\frac{\partial k}{\partial g}\frac{\partial g(\tau,q,e(\tau))}{\partial e} = C'(e(\tau)) && \textit{when } \tau > \tau_0 \end{aligned} \tag{5}$$

where τ^* is given implicitly by $[(1-F_P(\tau^*))\lambda + (1-F_U(\tau^*))(1-\lambda)]n = \bar{A}$ and τ_0 is given implicitly by $g(\tau_0,q,e) - g^* = \frac{\partial k}{\partial g}\frac{\partial g(\tau_0,q,e)}{\partial e} - C'(e) = 0$

(b) P students are overrepresented at university.

(c) Average realized human capital is higher for the P group.

Proof. See Appendix. □

This is a simultaneous move game – students choose effort levels knowing that everyone's choice affects everyone else's payoff via the threshold grade. However, the standard solution method is not feasible because payoffs are discontinuous and the point of discontinuity depends on the strategic choice variable. I therefore solve for equilibrium through the following novel solution method: (i) given a threshold grade g^* I find the optimal effort as a function of τ conditional on not attending university, (ii) I do the same conditional on attending university, (iii) I compare payoffs associated with both schedules to find the one which gives a higher payoff for each τ, and (iv) I establish the value of g^* consistent with all university seats being filled. This method separates the two decisions of whether to go to university, and the optimal choice of effort conditional on the first choice.

Figure 4 in the $e-\tau$ space depicts steps (i) and (ii) above. The green GG line is the locus of $e-\tau$ combinations that achieve the threshold grade g^*. It slopes downward since $g(\tau,q,e)$ is increasing in all its arguments. Below and to the left of this line, the student does not achieve the threshold grade and does not attend university. At all other points, the student achieves or exceeds the threshold grade and attends university. The orange $0N$ line corresponds to the first order condition associated with maximizing realized human capital conditional on not attending university. It is upward sloping since $g(\tau,q,e)$ has positive cross derivatives, so a higher τ means that the marginal returns to effort are higher. Similarly, the blue $0Y$ line depicts the effort level that maximizes realized human capital conditional on attending university, and is positively sloped. Moreover, $0Y$ lies above $0N$ because attending university increases human capital, so the marginal gain from increasing effort is higher.

Consider first the optimal effort schedule conditional on not attending university $e_N(\tau)$. This restricts the student to below GG . At levels of τ up to τ_1, $0N$ gives the optimal choice. After τ_1, effort levels consistent with $0N$ would put the student in university, so to ensure not attending university, effort levels must follow GG. Thus, $e_N(\tau)$ is the lower envelope of $0N$ and GG. Similarly, consider the optimal effort schedule conditional on attending university $e_Y(\tau)$. The student must be restricted to GG or above. For

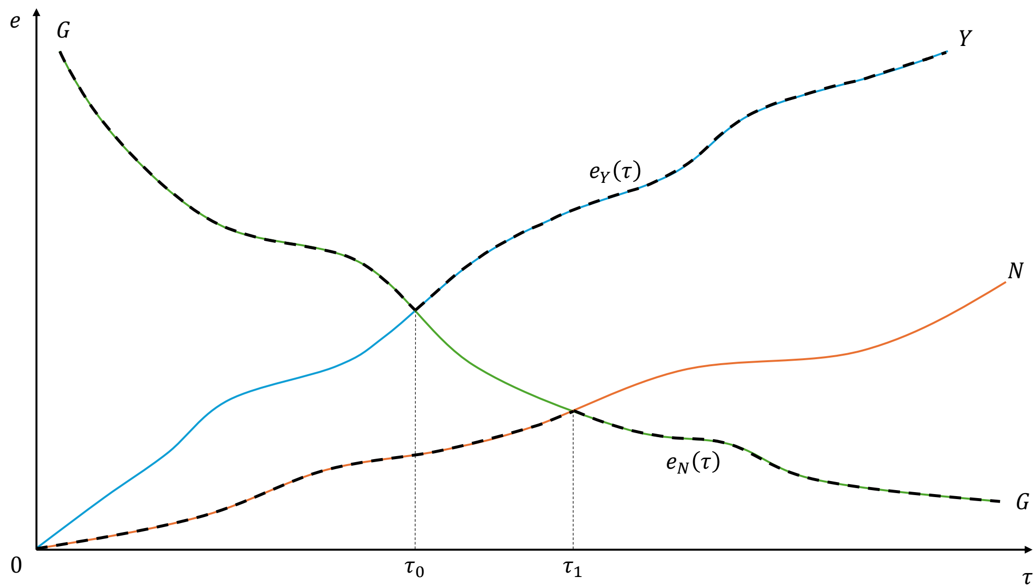

Figure 4: Optimal Effort Conditional on Attending and Not Attending University

low levels of τ up to τ_0 effort levels consistent with $0Y$ would not achieve the threshold grade, so effort levels must follow GG to ensure attending university. Above τ_0, $0Y$ gives optimal effort. $e_Y(\tau)$ is thus the upper envelope of GG and $0Y$.

Figure 5 depicts step (iii). $S_N(\tau)$ and $S_Y(\tau)$ show the payoffs corresponding to $e_N(\tau)$ and $e_Y(\tau)$ respectively. The Appendix shows that both are continuous and positively sloped. $S_N(0) = 0$ since the optimal effort is 0, but $lim_{\tau \to 0} S_Y(\tau) = -\infty$ since an arbitrarily low τ will require an arbitrarily high e to make the threshold grade, resulting in arbitrarily high costs. Moreover, $S_N(\tau_0) < S_Y(\tau_0)$. Unconstrained optimal effort levels conditional on attending university (depicted by $0Y$ in Figure 4) will always result in a higher payoff than the unconstrained optimal effort level conditional on not attending university (depicted by $0N$) because $k(g) > g$. $e_N(\tau_0)$ corresponds to a constrained effort level, which results in a lower payoff than the unconstrained effort level, which in turn is lower than the payoff from $e_Y(\tau_0)$ that corresponds to the unconstrained optimal effort level. Thus, $S_N(\tau)$ and $S_Y(\tau)$ must cross between 0 and τ_0. The Appendix shows they must cross only once. Denote this point τ^*. The highest payoff $S(\tau)$ is given by not attending university up to τ_* and attending university thereafter.

Figure 6 shows the resulting optimal effort schedule $e(\tau)$. Intuitively, for students with extremely low levels of τ (below τ^*) it takes too much effort to meet the threshold grade and the disutility of putting in that effort outweighs the extra returns. So they choose to slack and not attend university. Students with extremely high levels of τ (over τ_0) meet the threshold simply by maximizing their realized human capital. Students with moderate τ (between τ^* and τ_0) need to exert extra effort to meet the threshold grade, but

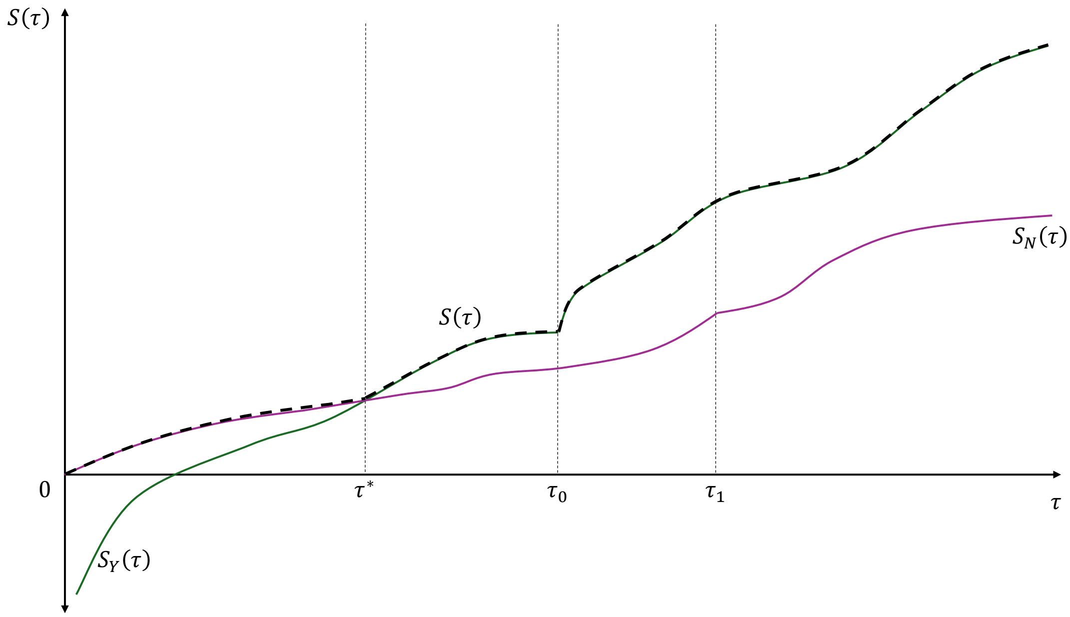

Figure 5: Comparing Payoffs from Attending and Not Attending University

the extra benefits of attending university outweigh the extra costs.

Step (iv) completes equilibrium determination. All students with $\tau \geq \tau^*$ attend university, so we must have $[(1 - F_P(\tau^*))\lambda + (1 - F_U(\tau^*))(1 - \lambda)]n = \bar{A}$. Since τ^* is determined by $S_N(\tau^*) = S_Y(\tau^*)$, and since $S_Y(\tau)$ in the relevant range depends on g^*, we only need to find a g^* corresponding to the τ^* that exactly clears college seats.

Since parental human capital enters τ, the τ distribution for P students dominates that for U students, resulting in overrepresentation at university. Average realized human capital is higher for the P group owing both to the overrepresentation at university and the higher average τ endowment.

4.3 Imposing Affirmative Action

Lemma 4.3. *Compared to the outcome with no affirmative action,*

(a) τ_0, τ_1 *and* τ^* *increase if the threshold grade increases and vice versa.*

(b) The threshold grade increases for P *students and reduces for* U *students.*

(c) In both groups, some students increase effort while others reduce effort. The effect on average effort is ambiguous.

(d) The effect on average realized human capital for both groups, and the difference between groups, is ambiguous.

Proof. See Appendix. □

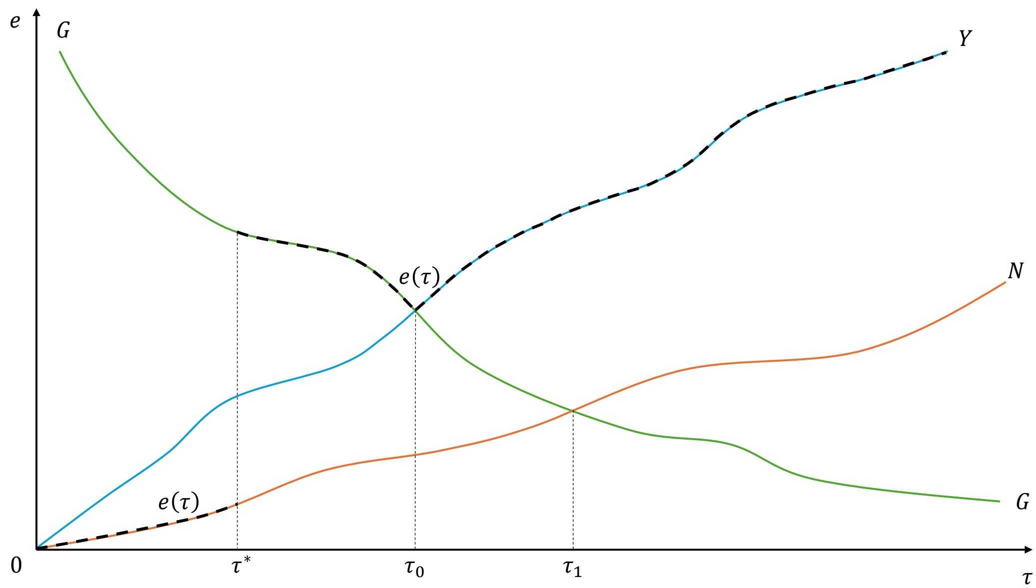

Figure 6: Optimal Effort

One can solve for equilibrium as earlier. In steps (i) and (ii), the change in threshold grade only affects GG. Suppose the threshold grade changes to $g_i^* > g^*$. Then Figure 7 depicts the consequent changes. GG shifts out and to the right to G_iG_i. Since the threshold grade is higher, for any given τ meeting the threshold grade requires more effort, and vice versa. Since G_iG_i is negatively sloped, it intersects $0N$ and $0Y$ to the right of where GG intersects them, resulting in $\tau_{0i} > \tau_0$ and $\tau_{1i} > \tau_1$. Moreover, the effort required conditional on attending university rises up to τ_{0i} and that conditional on not attending university rises beyond τ_1. The analysis is reversed for $g_i^* < g^*$, with $\tau_{0i} < \tau_0$, $\tau_{1i} < \tau_1$, effort conditional on attending university falling up to τ_0 and that conditional on not attending falling beyond τ_{1i}.

Figure 8 carries forward the analysis for $g_i^* > g^*$ to comparing payoffs and finding τ^*. Since effort conditional on attending university rises up to τ_i^*, it moves further away from the unconstrained optimum leading to a fall in $S_Y(\tau)$ in this range. Since effort conditional on not attending university falls beyond τ_1, it moves closer to or exactly to the unconstrained optimum, leading to a rise in $S_N(\tau)$ in this range. Crucially, $S_N(\tau)$ does nto change up to τ_1. This means the intersection point $\tau_i^* > \tau^*$. The analysis is reversed for $g_i^* < g^*$, with $S_Y(\tau)$ rising up to τ_0 and $\tau_i^* < \tau^*$.

Since $g_i^* > g^* \Rightarrow \tau_i^* > \tau^*$ (and vice versa), and since affirmative action requires $\tau_P^* > \tau^* > \tau_U^*$ to achieve equal representation, this necessarily required $g_P^* > g^* > g_U^*$. Figure 9 shows the effect of an increase in threshold grade for P students on optimal effort. Between τ^* and τ_P^* effort falls. This corresponds to the set of P students who just made it to university by exerting extra effort without affirmative action, but now no longer find it worthwhile to exert even more effort to meet the increased threshold. They settle for not

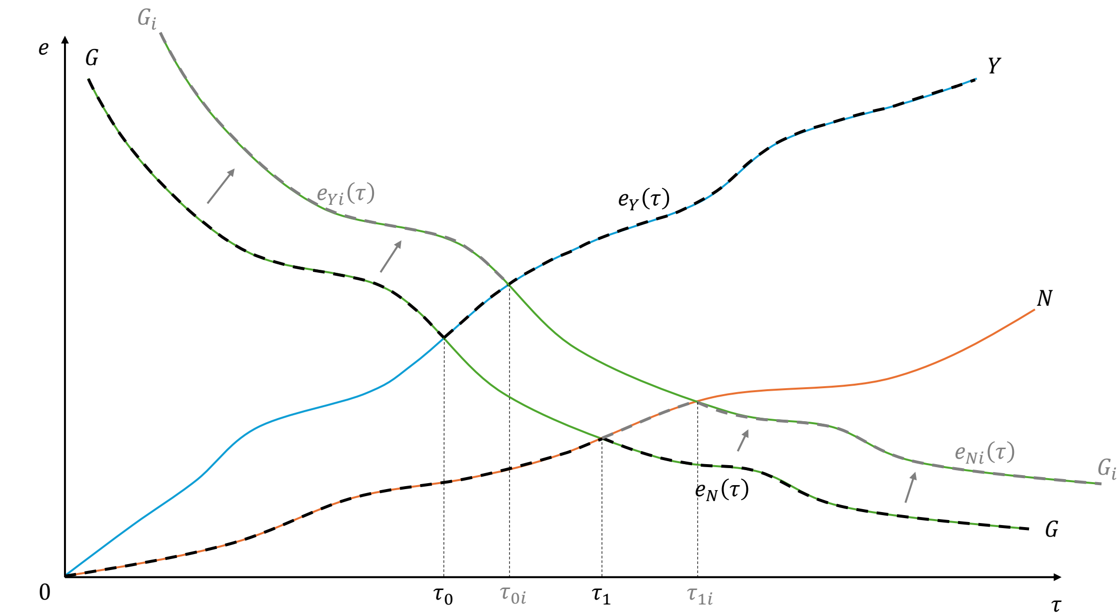

Figure 7: The Effect of Increasing g^* on τ_0 and τ_1

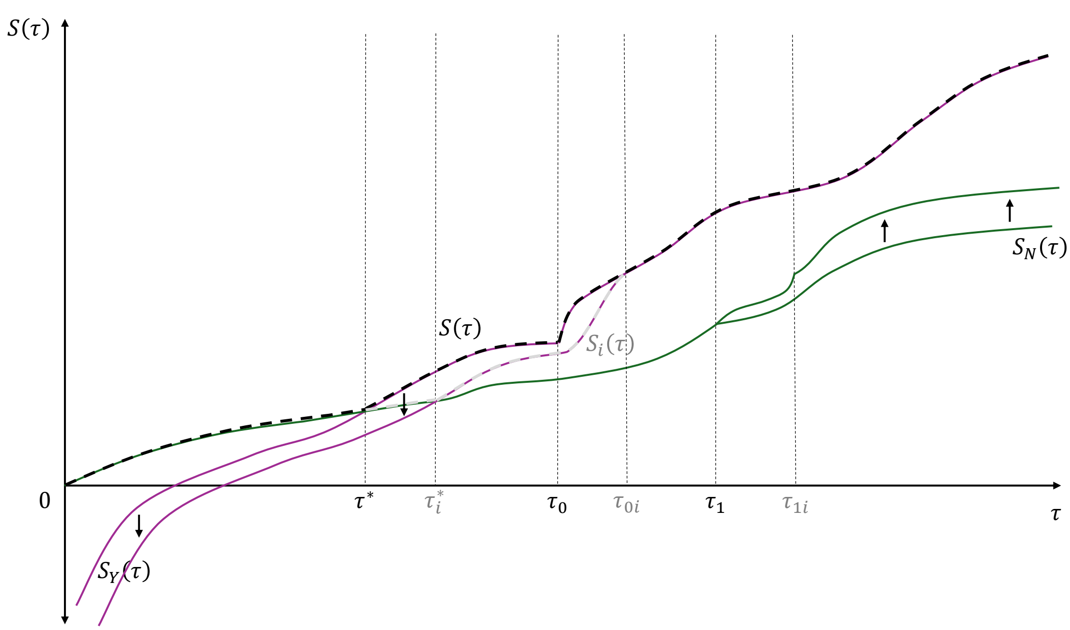

Figure 8: The Effect of Increasing g^* on τ^*

attending university reduce their effort level to the associated unconstrained optimum. Conversely, between τ_P^* and τ_{0P} effort rises. Students between τ_P^* and τ_0 needed to work extra hard to attain the threshold grade without affirmative action, and while they now need to put in even more effort, the gains from attending university are still greater than the extra costs. Students between τ_0 and τ_{0P}, who did not need to work extra hard to achieve the threshold grade without affirmative action, suddenly find they need to do so now, but putting in the extra work and attending university is still preferable to not attending university. Thus, the effect on average effort levels is ambiguous, and depends on the density of students in the relevant intervals.

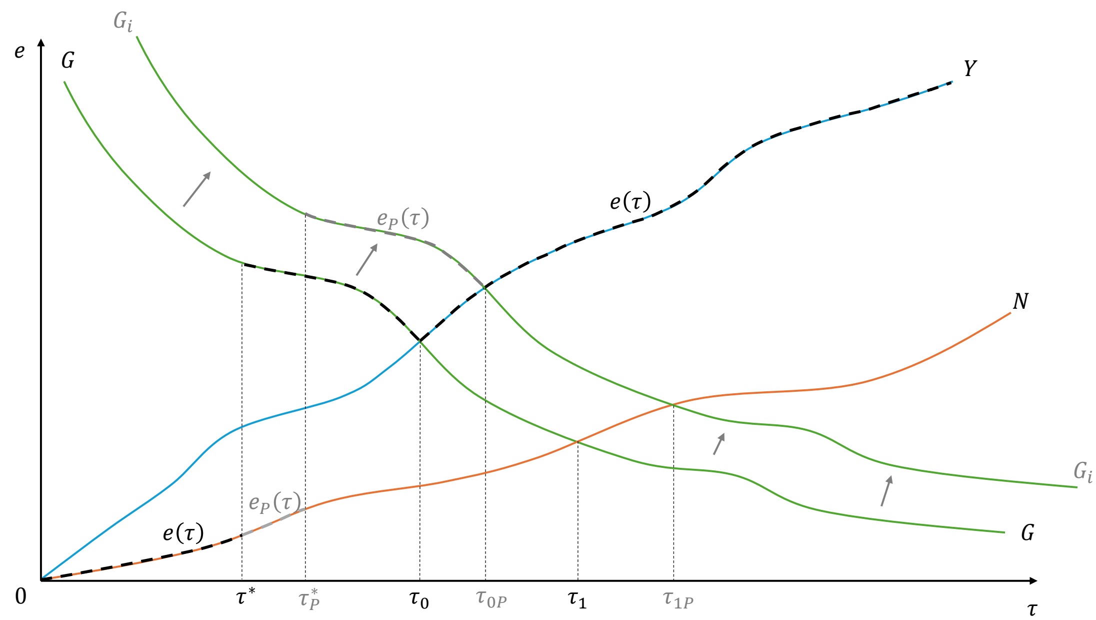

Figure 9: The Effect of Increasing g^* on Optimal Effort for P students

The analysis is reversed for U students, for whom $g_U^* < g^*$. Effort rises for students between τ_U^* and τ^*. These are students who earlier did not find it worthwhile to exert extra effort to achieve the threshold grade, but find it worthwhile after the threshold grade falls. Effort falls for students between τ^* and τ_0. These students were earlier working extra hard to meet the threshold grade, but now need to work less hard to meet the lower threshold grade. Students between τ_{0U} and τ_0 are able to shift to their unconstrained effort level. The effect on average effort is, again, ambiguous.

Since the effect on effort is ambiguous for both groups, the effect on average human capital for either group, and the difference between group averages, is ambiguous. This leads to the following proposition.

Proposition 4.1. *Affirmative action may cause average effort to rise for P students and fall for U students, which in turn might increase the difference between average realized human capital across groups.*

Figure 10 illustrates such a situation. The increase in effort for the U group in the range $[\tau_U^*, \tau^*)$ will likely be outweighed by the decline in the range $[\tau^*, \tau_0)$ to an extent that outweighs the effect of increased representation. Similarly, for P students, the reduction in effort in the range $[\tau^*, \tau_P^*)$ will likely be outweighed by an increase in effort in the range $[\tau_P^*, \tau_{0P})$ to an extent that outweighs the effect of reduced representation.

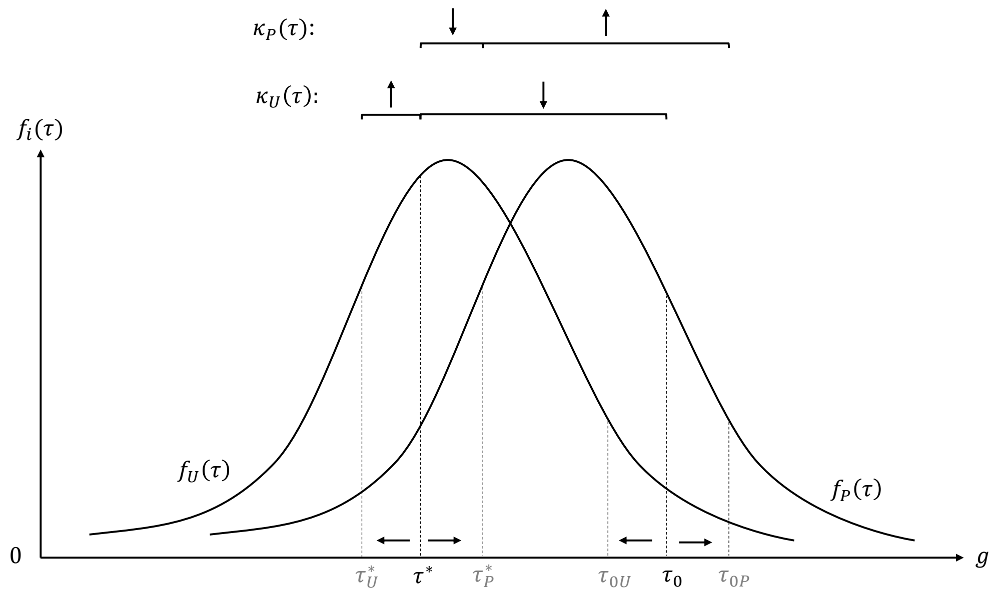

Figure 10: An Illustration of Increasing Inequality

5 Conclusion

This paper examines the effectiveness of affirmative action policies in higher education in reducing inter-group inequality when strategic responses by schools and students are taken into account. Students in the model are endowed with ability and parental human capital. The privileged group has a parental human capital distribution that dominates that of the underprivileged group, while ability is distributed similarly across groups. Students undergo schooling and may attend university if they achieve a threshold grade or greater. Both school and university raise realized human capital.

When school quality and student effort are fixed and homogeneous, privileged students perform better at school on account of having richer parents on average, and are overrepresented at university. Affirmative action targeted at achieving equal rates of representation across groups causes threshold grades to fall for the underprivileged group and rise for the privileged group. While the difference in average realized human capital across groups falls, it remains positive because privileged students still benefit from having richer parents on average.

When schools are able to choose quality sequentially, they compete to attract privileged group students. Early movers invest in high quality and educate privileged students while late movers invest in low quality and educate underprivileged group students. The disparity between privileged and underprivileged group average realized human capital is even wider than the baseline case because privileged students have higher parental capital on average, go to better quality schools and are overrepresented at university. Affirmative action makes privileged group students less attractive and underprivileged group students more attractive. This reduces the intensity of competition for privileged group students, leading to lower investment in quality by early movers, who still educate privileged group students. Investment in quality by late movers may also reduce if doing so puts a high density of students in the vicinity of the threshold grade, so that the policy pushes as many students into university as possible. Inequality between groups might increase as a result.

When students are able to choose effort levels, three groups emerge. Those with low levels of endowment choose to exert low effort levels and not attend university. Those with high levels of endowment attend university but exert moderate levels of effort since they comfortably exceed the threshold grade. Those with moderate levels of endowment would like to exert moderate levels of effort if they were guaranteed a place in university, but competition from other students means they must exert extra effort to achieve the threshold grade. Affirmative action causes effort to rise for some members and fall for other members of both groups. For the privileged group, a rise in the threshold grade means that it is no longer feasible to compete to attend university for some students with moderate endowment, who reduce their effort levels and choose not to attend university. However, others with moderate endowment increase their effort level in order to meet the higher threshold grade. For the underprivileged group, the fall in the threshold grade makes it worthwhile for some students with moderate endowment to work extra hard to meet the threshold grade and attend university where they earlier would not. However, the fall in threshold grade requires lower effort from those with moderate endowment who were earlier exerting extra to achieve the threshold. The net effect on average effort levels and average realized human capital is ambiguous for both groups, and inequality may increase as a result.

This paper shows the importance of considering strategic decision making by different actors when evaluating the effectiveness of affirmative action policies. It suggests possible reasons to investigate when affirmative action policies do not have the desired effect.

Extensions of this paper might include numerical simulations to understand dynamic effects when affirmative action policies are in place across generations. Empirical work might also use these theoretical models as bases for structural econometric models to evaluate the extent of strategic reactions affirmative action policies.

References

Arcidiacono, Peter (2005) "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?" *Econometrica*, 73 (5), 1477–1524.

Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan (2010) "Affirmative action in education: Evidence from engineering college admissions in India," *Journal of Public Economics*, 94 (1-2), 16–29.

De Fraja, Gianni (2002) "Affirmative action and efficiency in education," *Available at SSRN 315679*.

Durlauf, Steven N (2008) "Affirmative action, meritocracy, and efficiency," *Politics, Philosophy & Economics*, 7 (2), 131–158.

Epple, Dennis, Richard Romano, and Holger Sieg (2008) "Diversity and affirmative action in higher education," *Journal of Public Economic Theory*, 10 (4), 475–501.

Fryer Jr, Roland G, Glenn C Loury, and Tolga Yuret (2008) "An economic analysis of color-blind affirmative action," *The Journal of Law, Economics, & Organization*, 24 (2), 319–355.

Milgrom, P. and C. Shannon (1994) "Monotone comparative statics," *Econometrica*, 62 (1), 157–180.

del Rey Canteli, Elena and N Racionero (2008) "An efficiency argument for affirmative action in higher education," in *XV Encuentro de Economía Pública: políticas públicas y migración*, 44.

Rotthoff, Kurt W (2008) "Could affirmative action be efficient in higher education?" *Economics Letters*, 99 (3), 574–576.

A Appendix

Proof of Lemma 2.1.

(a) With q and e fixed, equation 1 implicitly defines p as a function of a parametrized by g. Call this function $P(a|g)$. Since grades increase in each argument (Assumption 2.2(ii)), if a increased then p would have to fall to achieve the same g. Hence $P(a|g)$ is a decreasing function for a fixed g. Moreover, since $a, p \geq 0$, this function exists in the positive quadrant. Further, curves corresponding to higher values of g lie strictly above and to the right of those corresponding to lower values.

The probability that a grade is below $\tilde{g}$ is then the probability mass of $H_i(a,p)$ that lies to the left and below $P(a|\tilde{g})$ plotted in the $p-a$ space:

$$\begin{aligned}
F_i(\tilde{g}) &= \int_a \int_0^{P(a|\tilde{g})} h_i(a,p)dpda \\
&= \int_a \int_0^{P(a|\tilde{g})} x_i(a)b_i(p|a)dpda \\
&= \int_a x_i(a) \int_0^{P(a|\tilde{g})} b_i(p|a)dpda \\
&= \int_a x_i(a)B_i(P(a|\tilde{g}))da \\
&= \mathbb{E}_a\left[B_i(P(a|\tilde{g}))\right]
\end{aligned}$$

(b) By Assumption 2.1(ii), $B_P(P(a|g)) < B_U(P(a|g))$. Since the marginal distribution of a is the same across groups (Assumption 2.1(i)), $\mathbb{E}_a\left[B_P(P(a|g))\right] < \mathbb{E}_a\left[B_U(P(a|g))\right]$.

(c) Follows from equating the sum of the numbers of students of both groups who get grades above g^* with the number of university seats $\bar{A}$.

(d) From (b) above, $1 - F_P(g^*) \geq 1 - F_U(g^*)$.

(e) Let $\bar{g} = g(\bar{a}, \bar{p}, q, e)$ be the highest attainable grade. For group i the average human capital given a threshold grade g^* is

$$\tilde{k}_i = \int_0^{g^*} gdF_i(g) + \int_{g^*}^{\bar{g}} k(g)dF_i(g) = \int_0^{\bar{g}} gdF_i(g) + \int_{g^*}^{\bar{g}} (k(g)-g)dF_i(g) = \mathbb{E}_{F_i}\left[g\right] + \int_{g^*}^{\bar{g}} (k(g)-g)dF_i(g)$$

Since $F_P(g) \leq F_U(g)$, $\mathbb{E}_{F_P}\left[g\right] \geq \mathbb{E}_{F_U}\left[g\right]$. Integrating the second term by parts,

$$\int_{g^*}^{\bar{g}} (k(g)-g)dF_i(g) = k(\bar{g}) - \bar{g} - \left(k(g^*) - g^*\right)F_i(g^*) - \int (k'(g)-1)F_i(g)dg$$

The result follows from Assumption 2.3.

□

Proof of Lemma 2.2.

(a) Equal representation requires separate threshold grades for each group g_P^* and g_U^* which satisfy the following conditions

$$1 - F_P(g_P^*) = 1 - F_U(g_U^*)$$
$$[(1 - F_P(g_P^*))\lambda + (1 - F_U(g_U^*))(1 - \lambda)]n = \bar{A}$$

Without affirmative action, from Lemma 2.1(b), $1 - F_P(g^*) \geq 1 - F_U(g^*)$. Moving from Lemma 2.1(c) to the first equation above while obeying the second equation requires $1 - F_P(g^*) \geq 1 - F_P(g_P^*)$ and $1 - F_U(g^*) \leq 1 - F_U(g_U^*)$. This requires $g_U^* \leq g^* \leq g_P^*$.

(b) Average realized human capital for group i is given by

$$\tilde{k}_i^{AA} = \mathbb{E}_{F_i}[g] + \int_{g_i^*}^{\bar{g}} (k(g) - g) dF_i(g)$$

Consequently,

$$\tilde{k}_P^{AA} - \tilde{k}_P = -\int_{g^*}^{g_P^*} (k(g) - g) dF_P(g) \leq 0$$
$$\tilde{k}_U^{AA} - \tilde{k}_U = \int_{g_U^*}^{g^*} (k(g) - g) dF_U(g) \geq 0$$

(c) The difference between average realized human capital for the P and U groups is

$$\tilde{k}_P^{AA} - \tilde{k}_U^{AA} = \underbrace{\mathbb{E}_{F_P}[g] - \mathbb{E}_{F_U}[g]}_{M} + \underbrace{\left[\int_{g_P^*}^{\bar{g}} (k(g) - g) dF_P(g) - \int_{g_U^*}^{\bar{g}} (k(g) - g) dF_U(g)\right]}_{N}$$

$M \geq 0$ by Lemma 2.1(b). Integrating N by parts and using $F_P(g_P^*) = F_U(g_U^*)$,

$$N = -\underbrace{F_U(g_U^*)[(k(g_P^*) - g_P^*) - (k(g_U^*) - g_U^*)]}_{N1}$$
$$+ \underbrace{\left[\int_{g_U^*}^{\bar{g}} (k'(g) - 1) F_U(g) dg - \int_{g_P^*}^{\bar{g}} (k'(g) - 1) F_P(g) dg\right]}_{N2}$$

By Lemma 2.1(b), since $F_U(g)$ is weakly increasing, and by Assumption 2.3(ii),

$$\begin{aligned}
N2 &> \int_{g_U^*}^{\bar{g}} (k'(g)-1)F_U(g)dg - \int_{g_P^*}^{\bar{g}} (k'(g)-1)F_U(g)dg \\
&= \int_{g_U^*}^{g_P^*} (k'(g)-1)F_U(g)dg \\
&\geq F_U(g_U^*) \int_{g_U^*}^{g_P^*} (k'(g)-1)dg \\
&= [k(g)-g]_{g_U^*}^{g_P^*} \\
&= N1
\end{aligned}$$

Hence, $N = -N1 + N2 > -N1 + N1 = 0$

(d) Economy-wide average realized human capital without affirmative action is $\tilde{k} = \lambda\tilde{k}_P + (1-\lambda)\tilde{k}_U$, and $\tilde{k}^{AA}$ is defined analogously for when affirmative action is imposed. Their difference is

$$\tilde{k}^{AA} - \tilde{k} = \int_{g_U^*}^{g^*} (1-\lambda)(k(g)-g)dF_U(g) - \int_{g^*}^{g_P^*} \lambda(k(g)-g)dF_P(g)$$

Affirmative action takes away university seats from P students with grades in the range $[g^*, g_P^*]$ and gives them to U students with grades in the range $[g_U^*, g^*]$. The size of both these groups must be the same, therefore

$$\begin{aligned}
(1-\lambda)n \int_{g_U^*}^{g^*} dF_U(g) &= \lambda n \int_{g^*}^{g_P^*} dF_P(g) \\
\Rightarrow \quad (k(g^*)-g^*) \int_{g_U^*}^{g^*} (1-\lambda)dF_U(g) &= (k(g^*)-g^*) \int_{g^*}^{g_P^*} \lambda dF_P(g)
\end{aligned}$$

Since $k(g)-g$ is weakly increasing in g by Assumption 2.3(ii), $(k(g^*)-g^*)\int_{g^*}^{g_P^*} \lambda dF_P(g) \leq \int_{g^*}^{g_P^*} \lambda(k(g)-g)dF_P(g)$ and $(k(g^*)-g^*)\int_{g_U^*}^{g^*}(1-\lambda)dF_U(g) \geq \int_{g_U^*}^{g^*}(1-\lambda)(k(g)-g)dF_U(g)$. Applying this to the above,

$$\begin{aligned}
\int_{g_U^*}^{g^*} (1-\lambda)(k(g)-g)dF_U(g) &\leq \int_{g^*}^{g_P^*} \lambda(k(g)-g)dF_P(g) \\
\Rightarrow \quad 0 &\leq \int_{g^*}^{g_P^*} \lambda(k(g)-g)dF_P(g) - \int_{g_U^*}^{g^*} (1-\lambda)(k(g)-g)dF_U(g) = \tilde{k}^{AA} - \tilde{k}
\end{aligned}$$

with equality when $k'(g) = 1$.

□

Proof of Lemma 3.1. Parts (a) and (b) are analogous to the proofs of Lemma 2.1(a) and (b), with the addition of p also being parametrized by q since q is now stochastic. For part (c), applying Assumption

2.2(ii), an increase in q requires a decrease in p to achieve the same g. Hence $q'' \geq q' \Rightarrow P(a|g,q'') \leq P(a|g,q') \Rightarrow F_i(a|g,q'') \leq F_i(a|g,q')$, where the last step uses Lemma 3.1(a).

For parts (d)-(f), I first find best response strategies using backwards induction relying on the Subgame Perfect Equilibrium concept.

Students' acceptance of offers. Students know (a,p) and e is constant, so their expected payoff is $\theta(q) = Prob(g(a,p,q,e) < g^*)g(a,p,q,e) + Prob(g(a,p,q,e) \geq g^*)k(g(a,p,q,e))$. Increasing q increases both $g(a,p,q,e)$ and $k(g(a,p,q,e))$, and shifts probability towards $k(g(a,p,q,e))$, all of which increase $\theta(q)$. Hence, students prefer higher quality schools but are indifferent between schools of the same quality.

Schools' offers. A schools' payoff is given by

$$\begin{aligned} S(q|i,g^*) &= Prob(g \geq g^*|i)\mathbb{E}\left[k(g)|i, g \geq g^*\right] + Prob(g < g^*|i)\mathbb{E}\left[g|i, g < g^*\right] - C(q) \\ &= (1 - F_i(g^*|q))\frac{\int_{g^*}^{\bar{g}} k(g)dF_i(g|q)}{1 - F_i(g^*|q)} + F_i(g^*|q)\frac{\int_0^{g^*} gdF_i(g|q)}{F_i(g^*|q)} - C(q) \\ &= \mathbb{E}_{F_i|q}\left[g\right] + \int_{g^*}^{\bar{g}}(k(g) - g)dF_i(g|q) - C(q) \end{aligned}$$

Following the method of the proof of Lemma 2.1(e), integrating the second term by parts and applying Lemma 3.1(c) gives $S(q|P,g^*) > S(q|U,g^*)$. Hence schools prefer P students to U students but are indifferent between students of the same group.

Schools' quality choice. Since all schools will make offers to P students first, and since P students prefer higher quality to lower quality schools, P students will be distributed among the λn highest quality schools. Hence, to attract a P student a school needs to be within the top λn quality schools.

Note that $S(q|i,g^*)$ is continuous since $g(a,p,q,e)$, $k(g)$, $X_i(a)$ and $B_i(p|a)$ are all continuous and $S(0|i,g^*) = 0$ since $g(a,p,0,e) = 0$ (Assumption 2.2(i), (iv)). By Assumption 3.1(iii)-(iv) $\arg\max_q S(q|i,g^*) \neq \phi$ and $S(q_B|U,g^*) > 0$. $q_G > q_B$ since $S(q|P,g^*) > S(q|U,g^*)$.

Consider the n-th school. There are two possibilities:

1. At least λn schools have chosen quality at least q_G. Choosing $q > q_G$ will attract a P student for sure but $S(q|P,g^*,q > q_G) < \max_q S(q|U,g^*)$. Choosing $q = q_G$ will attract a P student with some probability and give less expected payoff than $\max_q S(q|U,g^*)$. Choosing $q < q_G$ will attract a U student for sure, and in this range choosing q_B gives the highest payoff.
2. Fewer than λn schools have chosen quality at least q_G. Let q' be the minimum of the highest λn qualities chosen previously. $\{q \in (q', q_G] : S(q|P,g^*) \geq S(q_G|P,g^*)\} \neq \phi$ since $\max_q S(q|P,g^*) > \max_q S(q|U,g^*) = S(q_G|P,g^*)$ and $S(q|P,g^*)$ is continuous. Choosing $\arg\max_{q \in (q',q_G)} S(q|P,g^*)$ would attract a P student for sure and give payoff $\geq S(q_G|P,g^*)$. As earlier, choosing $q < q'$ would attract a U student for sure which would give at most $\max_q S(q|U,g^*) = S(q_G|P,g^*)$ and choosing $q > q_G$ would result in a lower payoff than choosing q_B.

Consider the $(n-1)$-th school. There are three possibilities.

1. At least λn schools have chosen quality at least q_G. As above, it should choose q_B.
2. Fewer than λn schools have chosen quality at least q_G. Let q' be the minimum of the top $\lambda n - 1$ qualities chosen so far.
   (a) $q' < q_G$. Choosing $\arg\max_{q \in (q', q_G)} S(q|P, g^*)$ would attract a P student for sure since even if the subsequent school chose q_G it would still be in the top λn. Moreover, it would get a payoff $\geq S(q_G|P, g^*)$.
   (b) $q' = q_G$. Choosing q_G would ensure that it would be a part of the top λn schools and attract a P student for sure since the subsequent school would choose q_B.

The analysis is analogous for the $(n-2)$-th school up to the $(1-\lambda)n$-th school. The best response of representative school k within this group can be summarized as: if at least λn schools have chosen at least q_G, choose q_B; if fewer than λn schools have chosen at least q_G, then if the minimum of the top $\lambda n - k$ qualities so far is less than q_G choose $\arg\max_{q \in (q', q_G)} S(q|P, g^*)$, else choose q_G.

For the λn-th and previous schools, the only way to ensure being in the top λn schools eventually is to choose q_G. Any $q < q_G$ would allow subsequent schools to choose higher quality and attract P students, and any $q > q_G$ gives a lower payoff than q_G.

All the best responses taken together imply the results in (d)-(f).

From parts (b) and (c) above, $1 - F_P(g^*|q_G) > 1 - F_U(g^*|q_G) \geq 1 - F_U(g^*|q_B)$, proving (g).

Analogous to the baseline model, average human capital for each group is given by

$$\tilde{k}_P = \mathbb{E}_{F_P|q_G}[g] + \int_{g^*}^{\bar{g}} (k(g) - g) dF_P(g|q_G)$$
$$\tilde{k}_U = \mathbb{E}_{F_U|q_B}[g] + \int_{g^*}^{\bar{g}} (k(g) - g) dF_U(g|q_B)$$

Integrating by parts and applying (b) and (c) proves (h). □

Proof of Lemma 3.2. To compute the new equilibrium consider how best response strategies change.

Students still prefer higher quality over lower quality schools. The proof is the same as in that for Lemma 3.1, with g^* replaced by g_i^*.

To analyze schools' preferences between groups, consider first the effect of a change in threshold grade on $S(q|i, g^*)$.

$$\frac{\partial S(q|i, g^*)}{\partial g^*} = -[k(g^*) - g^*] f_i(g^*|q) < 0$$

by Assumption 2.2(ii). Since $g_P^* > g^* > g_U^*$, $S(q|P, g_P^*) < S(q|P, g^*)$ and $S(q|U, g_U^*) > S(q|U, g^*)$, i.e. P students are less attractive and U students more attractive at a given quality level. However, $S(q|P, g_P^*) > S(q|U, g_U^*)$, i.e. P students remain more attractive than U students. To see this, note that

$$S(q|P, g_P^*) - S(q|U, g_U^*) = \mathbb{E}_{F_P|Q}[g] - \mathbb{E}_{F_U|Q}[g] + \int_{g_P^*}^{\bar{g}} (k(g) - g) dF_P(g|q) - \int_{g_U^*}^{\bar{g}} (k(g) - g) dF_U(g|q) > 0$$

following the same proof as for Lemma 2.2(c) replacing $F_i(g)$ with $F_i(g|q)$.

Since the new payoff curves for schools follow the same rules as without Affirmative Action, the analysis for quality choice remains the same as in that for Lemma 3.1, with q'_G and q'_B corresponding to the new payoffs.

Combined, these strategies lead to (a)-(c).

Now consider how quality choices change for good quality schools. Since $g^*_U < g^*$, $S(q|U, g^*_U) > S(q|U, g^*)$ everywhere and consequently $\max S(q|U, g^*_U) > \max S(q|U, g^*)$. Moreover, since $g^*_P > g^*$, $S(q|P, g^*_P) < S(q|P, g^*)$. Given continuity, this implies $q'_G = \max\{q : S(q|P, g^*_P) = \max S(q|U, g^*_U)\} < \max\{q : S(q|P, g^*) = \max S(q|U, g^*)\} = q_G$.

To analyze how bad quality schools' choice changes, I appeal to a version of a monotone comparative statics result from Milgrom and Shannon (1994) reproduced below as Lemma A.1. This application satisfies the conditions for Lemma A.1. The function being maximized is $f(x, t) = S(x|U, t)$, and $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$. $(\mathbb{R}_+, \geq)$, where '$\geq$' is the euclidean relation, is both a partially ordered set[9] and a lattice[10]. Any function where the choice set is a scalar is always quasisupermodular in the choice variable[11]. $\arg\max_q S(q|U, g)$ moves in the same direction as g if increasing differences, which is sufficient for single crossing differences from negative to positive[12] holds. Increasing differences implies

$$\frac{\partial^2 S(q|U,g)}{\partial q \partial g} \geq 0 \Rightarrow -(k(g) - g)\frac{\partial f_U(g|q)}{\partial q} \geq 0 \Rightarrow \frac{\partial f_U(g|q)}{\partial q} \leq 0$$

using Assumption 2.3. Since $g^*_U < g^*$, $q'_B \leq q_B$ if this condition holds and $q'_B \geq q_B$ otherwise. This completes the proof for (d).

The difference in average realized human capital is identical to $S(q|P, g^*_P) - S(q|U, g^*_U)$, which I showed above is positive. Given $g^*_P > g^*$ and $q'_G < q_G$, it is easy to verify average realized human capital for the P group falls. However, the change for the U group is ambiguous because the associated change in school quality is ambiguous. The effect on economy-wide average human capital is consequently also ambiguous. □

Lemma A.1. *Let X be a lattice, let T be a partially ordered set, and let the function $f : X \times T \to \mathbb{R}$ have the following properties:*

1. *$\forall t \in T$, $f(\cdot, t)$ is quasisupermodular in x.*
2. *$f(x, t)$ has single crossing differences (SCD) from negative to positive.*

[9] A set $(S, \geq)$ is a partially ordered set if $\forall x, y, z \in S$, (i) $x \geq x$ (reflexivity), (ii) $x \geq y$ & $y \geq x \Rightarrow x = y$ (anti-symmetry), and (iii) $x \geq y$ & $y \geq z \Rightarrow x \geq z$ (transitivity).

[10] A partially ordered set $(s, \geq)$ is a lattice if every two element subset of S has an infimum and a supremum.

[11] A function $f : X \to \mathbb{R}$ is QSM if $f(x) - f(\inf\{x, x'\}) \geq 0 \Rightarrow f(\sup\{x, x'\}) - f(x') \geq 0$.

[12] A function $f : X \times T \to \mathbb{R}$ has single crossing differences if whenever $x'' > x'$, the function $g(t) = f(x'', t) - f(x', t)$ crosses 0 only once from negative to positive. Increasing differences, or $g(t)$ being increasing in t, is sufficient for single crossing differences from negative to positive. It can also be shown that if $X \subseteq \mathbb{R}^l$ and $T \subseteq \mathbb{R}^k$, then $f(x, t)$ has ID if

$$\frac{\partial^2 f}{\partial x_i \partial t_j} \geq 0 \;\forall i = 1, \ldots, l, j = 1, \ldots, k.$$

The reversed inequality corresponds to decreasing differences, which is sufficient for single crossing differences from positive to negative.

Then $\arg\max_{x\in X} f(x,t'') \geq \arg\max_{x\in X} f(x,t')$ *if* $t'' > t'$*. The converse holds if the SCD is from positive to negative.*

Proof. See Milgrom and Shannon (1994). □

Proof of Lemma 4.1.

(a) $g(a,p,q,e) = g(\tau(a,p),q,e) \Rightarrow \frac{\partial g}{\partial a} = \frac{\partial g}{\partial \tau}\frac{\partial \tau}{\partial a}$. $\frac{\partial g}{\partial a} > 0$ from Assumption 2.2(ii) and $\frac{\partial \tau}{\partial a} > 0$ from Assumption 4.1(i), so $\frac{\partial g}{\partial \tau} > 0$.

$g(a,p,q,e) = g(\tau(a,p),q,e) \Rightarrow \frac{\partial^2 g}{\partial e \partial a} = \frac{\partial \tau}{\partial a}\frac{\partial^2 g}{\partial e \partial \tau}$. From Assumption 2.2(iii) $\frac{\partial^2 g}{\partial e \partial a} > 0$, so $\frac{\partial^2 g}{\partial e \partial \tau}$.

Similarly, by differentiating with respect to a and q I can prove that $\frac{\partial^2 g}{\partial \tau \partial q} > 0$; and by differentiating with respect to e and q I can prove that $\frac{\partial^2 g}{\partial e \partial q} > 0$.

(b) Analogous to the proof of Lemma 2.1(a) using $\tau(a,p)$ as the starting function instead of $g(a,p,q,e)$.

(c) Analogous to the proof of Lemma 2.1(b) replacing g with τ.

□

Proof of Lemma 4.2.

(a) I use the following steps to find the equilibrium: (i) given a threshold grade g^*, find the optimal effort schedule given the student does not attend university $e_N(\tau)$, (ii) given g^* find optimal effort schedule given the student attends university $e_N(\tau)$, (iii) compare payoffs associated with both schedules to find the one that gives a higher payoff at each τ, (iv) find the value of g^* consistent with university seats being filled.

Step 1. By Lemma 4.1 $g(\tau,q,e) = g^*$ corresponds to a decreasing function in the $e - \tau$ space parameterized by g^*. Denote this function $e_G(\tau)$. $S(e|\tau, \text{not attend university}) = g(\tau,q,e) - C(e)$, with the first order condition $\frac{\partial g(\tau,q,e)}{\partial e} = C'(e)$. Denote the implicit relationship between e and τ defined by the first order condition as $e_N^u(\tau)$. Implicitly differentiating the first order condition gives $\frac{\partial e}{\partial \tau} = \left(C''(e) - \frac{\partial^2 g}{\partial e^2}\right)^{-1} \frac{\partial^2 g}{\partial e \partial \tau} > 0$ by Lemma 4.1(a) and Assumption 4.1(c). $e_N^u(0) = 0$ since $g(0,q,e) = 0$ and maximizing the payoff is then equivalent to minimizing $C(e)$. Let $e_G(\tau)$ and $e_N^u(\tau)$ intersect at τ_1. The payoff is concave in e since the second derivative $\frac{\partial^2 g}{\partial \tau^2} - C''(e) < 0$ by Assumption 4.1(iii). Then

$$e_N(\tau) = \begin{cases} e_N^u(\tau) & \text{if } \tau < \tau_1 \\ e_G(\tau) & \text{if } \tau \geq \tau_1 \end{cases}$$

Step 2. $S(e|\tau, \text{attend university}) = k(g(\tau,q,e)) - C(e)$, with the first order condition $\frac{\partial k}{\partial g}\frac{\partial g(\tau,q,e)}{\partial e} = C'(e)$. Denote the implicit relationship between e and τ defined by the first order condition as $e_Y^u(\tau)$. $e_Y^u(\tau) > e_N^u(\tau)$ becausae $k(g) > g$. Implicitly differentiating the first order condition, $\frac{\partial e}{\partial \tau} =$

$\left(C''(e) - \frac{\partial k}{\partial g}\frac{\partial^2 g}{\partial e^2} - \left(\frac{\partial g}{\partial e}\right)^2 \frac{\partial^2 k}{\partial g^2}\right)^{-1} \left(\frac{\partial k}{\partial g}\frac{\partial^2 g}{\partial \tau \partial e} + \frac{\partial^2 k}{\partial g^2}\frac{\partial g}{\partial e}\frac{\partial g}{\partial \tau}\right) > 0$ by Assumption 4.1 and Lemma 4.1(a). ${}^u_Y(0) = 0$ since $k(g(0, q, e)) = 0$ and maximizing the payoff is then equivalent to minimizing $C(e)$. Let $e_G(\tau)$ and $e^u_Y(\tau)$ intersect at $\tau_0 < \tau_1$. Then

$$e_Y(\tau) = \begin{cases} e_G(\tau) & \text{if } \tau \leq \tau_0 \\ e^u_Y(\tau) & \text{if } \tau > \tau_0 \end{cases}$$

Step 3. The realized payoff conditional on not attending and attending university are

$$S_N(\tau) \equiv S(e_N(\tau)|\tau, \text{not attend university}) = g(\tau, q, e_N(\tau)) - C(e_N(\tau))$$
$$S_Y(\tau) \equiv S(e_Y(\tau)|\tau, \text{attend university}) = k(g(\tau, q, e_Y(\tau))) - C(e_Y(\tau))$$

$S_N(0) = 0$ since $e_N(0) = 0$ and $lim_{\tau \to 0^+} S_Y(\tau) = -\infty$ since for any $g^* > 0$ an arbitrarily small τ would need an arbitrarily large e. It is easy to verify that $S'_N(\tau) > 0$ and $S'_Y(\tau) > 0$ everywhere. Both curves are continuous because $k(g)$, $g(\tau, q, e)$, $e_N(\tau)$ and $e_Y(\tau)$ are continuous. Moreover $S_Y(\tau) > S_N(\tau)\ \forall \tau \geq \tau_0$. For $\tau_0 \leq \tau \leq \tau_1$ both curves correspond to the unconstrained maximum, and $\max_e\{k(g(\tau, q, e)) - C(e)\} > \max_e\{g(\tau, q, e) - C(e)\}$ since $k(g) > g$. For $\tau > \tau_1$ $S_N(\tau)$ corresponds to the constrained maximum, which by definition gives a lower payoff than the unconstrained maximum, which is always lower for $S_N(\tau)$. Lastly, for $\tau < \tau_0$, $e_Y(\tau) > e^u_Y(\tau)$, which requires $\frac{\partial g(\tau, q e(e_Y(\tau)))}{\partial e} - C'(e_Y(\tau)) < 0$. This further implies that in this range $S'_Y(\tau) = \left(\frac{\partial g}{\partial e}\right)^{-1} C'(e)\frac{\partial g}{\partial \tau} > \frac{\partial g}{\partial \tau} = S'_N(\tau)$. Thus, $\exists \tau^*$ such that $S_N(\tau^*) = S_Y(\tau^*)$. Optimal effort is thus

$$e(\tau) = \begin{cases} e_N(\tau) & \text{when } \tau \leq \tau^* \\ e_Y(\tau) & \text{when } \tau > \tau^* \end{cases}$$

Step 4. τ must satisfy $[(1 - F_P(\tau^*))\lambda + (1 - F_U(\tau^*))(1 - \lambda)]n = \bar{A}$ to ensure all university seats are filled. Since τ^* is determined by $S_N(\tau^*) = S_Y(\tau^*)$ in the range where $S_Y(\tau)$ depends on g^*, it is possible to find the g^* which corresponds to the τ^* that just clears college seats.

(b) From Lemma 4.1(c), $1 - F_P(\tau^*) > 1 - F_U(\tau^*)$.

(c) Realized human capital is given by $\kappa(\tau) \equiv S(e(\tau)|\tau)$. Average realized human capital for group i is $k_i = \int \kappa(\tau) dF_i(\tau) = \int \kappa'(\tau)(1 - F_i(\tau))d\tau$, where the last equality flows from integration by parts. $\kappa'(\tau) > 0$ since all its constituent parts have positive slope. The result follows from part (b).

□

Proof of Lemma 4.3.

(a) Let $g_i^* > g^*$ and let $e_{Gi}(\tau)$ be defined implicitly by $g(\tau, q, e_{Gi}(\tau)) = g_i^*$. $g_i^* > g^* \Rightarrow e_{Gi}(\tau) > e_G(\tau)$ by Lemma 4.1(a). $e_N^u(\tau)$ and $e_Y^u(\tau)$ are unchanged. Let τ_{0i} and τ_{1i} be defined by $e_{Gi}(\tau_{0i}) = e_Y^u(\tau_{0i})$ and $e_{Gi}(\tau_{1i}) = e_N^u(\tau_{1i})$ respectively. Since $e'_{Gi}(\tau) < 0$, $e_N^{u'}(\tau) > 0$ and $e_Y^u(\tau) > 0$, $\tau_{0i} > \tau_0$ and $\tau_{1i} > \tau_1$. Following the proof for Lemma 4.2(a), $e_{Yi}(\tau) = e_{Gi}(\tau) > e_Y(\tau)$ for $\tau < \tau_{0i}$. Since e has moved further from the unconstrained optimal, $S_{Yi}(\tau) < S_Y(\tau)$ for $\tau < \tau_{0i}$, implying $\tau_i^* > \tau^*$. The analysis is reversed for $g_i^* < g^*$.

(b) Analogous to that of Lemma 2.2 replacing g^*, g_P^* and g_U^* with τ^*, τ_P^* and τ_U^*.

(c) The new effort schedule $e_i(\tau)$ is found in the same manner as in Lemma 4.2(a). Since $e_{GP}(\tau) > e_G(\tau)$, it is easy to verify that $e_P(\tau) < e(\tau)$ for $\tau^* \le \tau \le \tau_P^*$ and $e_P(\tau) > e(\tau)$ for $\tau_P^* < \tau < \tau_{0P}$. Since $e_{GU}(\tau) > e_G(\tau)$, it is easy to verify that $e_U(\tau) > e(\tau)$ for $\tau_U^* \le \tau \le \tau^*$ and $e_U(\tau) < e(\tau)$ for $\tau^* < \tau < \tau_0$. The effect on average effort depends on the relative density of students who increase and decrease effort.

(d) Denote the new realized human capital for group i student with endowment τ by $\kappa_i(\tau)$. It is easy to verify that $\kappa_U(\tau) > \kappa(\tau)$ for $\tau_U^* \le \tau \le \tau^*$, $\kappa_U(\tau) < \kappa(\tau)$ for $\tau^* < \tau < \tau_0$, $\kappa_P(\tau) < \kappa(\tau)$ for $\tau^* \le \tau \le \tau_P^*$, and $\kappa_P(\tau) >< \kappa(\tau)$ for $\tau_P^* < le\tau < \tau_{0P}$. The effect on average realized human capital for each group, and the difference between group averages, depends on the density of students in these intervals.

□