UNIVERSITÀ DEGLI STUDI FIRENZE

DISEI
DIPARTIMENTO DI SCIENZE PER L'ECONOMIA E L'IMPRESA

Working Papers - Economics

University Admission: Is Achievement a Sufficient Criterion?

Alessandro Tampieri

Working Paper N. 26/2019

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

University Admission: Is Achievement a Sufficient Criterion?*

Alessandro Tampieri†
University of Florence

September 1, 2019

Abstract

We analyse university admission through a statistical discrimination model where students differ in ability and social groups. Universities aim to enrol the students with the best human capital, which is given by their innate ability and of the learning carried out while at school. Students and school choose their learning and teaching effort based on the behaviour of universities. Interestingly, we find that students from less advantaged groups need a lower grade to be admitted to the best universities, while less competitive universities do the opposite. If a university cannot discriminate according to social groups, all students with same grade will attend universities of the same quality, with different levels of human capital according to their social group.

JEL Numbers: I21, I23, J71
Keywords: discrimination, affirmative action, studying effort, teaching effort.

*I am deeply indebted with Gianni De Fraja for many discussions that lead to substantial improvements of the paper. All errors are my own.

†Department of Economics and Business, University of Florence, email: alessandro.tampieri@unifi.it

1 Introduction

Should universities select which students among their applicants to admit exclusively on the basis of variables that measure the applicant's achievement, such as exams and admission tests, or should information on characteristics which are not relevant for the job, such as race, sex, the secondary school attended or the family social background also influence this choice? Most observers would be inclined to view a university choosing a student applicant in preference of a better qualified one of a different sex or age or race or "poshness" as practicing blatant discrimination. And yet such practice is used by many leading universities. In the US, the debate hinges around favourable treatment of racial minorities, typically blacks.[1] In the UK the focus is instead on the social background and the kind of school, private or state funded, attended by the applicants.[2]

Those who favour basing admissions on more than just past academic achievement, argue that candidates with different backgrounds, bring benefits (to the institution, or to the wider society) that exceed the cost of excluding more qualified candidates. In this paper we show instead that different admission standards for different groups would be chosen by a university with no other objective than admitting the best candidates. This *prima facie* surprising result is in fact a natural consequence of the well understood (since at least Spence 1973's seminal paper) dual role of education, which both enhances human capital and signals an individual's innate ability.

We construct a model of statistical discrimination where students attend school and then take a final school test, which determines their university career. Students differ in ability, which affects the variable of interest to universities directly, and in other characteristics – sex, race, socio-economic background, the type of school they attended – which instead do not. Universities aim to enrol the students with the best human capital: this is a combination of the innate ability and of the learning carried out while at school. Universities know a student's characteristics and observe the result of her test. Based on this information universities update their belief of the student's human capital, according to Bayes rule. Students and school understand the behaviour of universities and adjust their learning and teaching effort accordingly.

The main result of our paper is that students from groups with a nosier signal need a lower grade to be admitted to the best universities. To the extent that less advantaged groups of students have indeed a nosier distribution of abilities, this is the opposite of the established

[1] Pacelli (2011) reviews the four US Supreme Court landmark cases – Bakke, Hopwood, Grutter, and Parents Involved – in her analysis of the latest such case to appear before the Supreme Court, Abigail Fisher's claim that she was denied admission by University of Texas Austin, while less qualified Latinos and African-Americans were offered places. The issue got media coverage in the (New York Times (2012a), (2012b) and (2016), *inter alia*.

[2] See DfES, 2003. Several other countries in the world impose some form of affirimative action (Sowell 2004). Most recently, Brazil passed a law reserving half of the university places to students from state schools and increasing the number of students of African descent.

model of statistical discrimination in the labour market (along the lines of Phelps 1972 and Lundberg and Startz 1983), in that in our model they appear to be discriminated favourably. In fact universities are risk neutral and simply maximise the expected human capital of their intake. Indeed we show that less competitive universities impose stricter admission standard on students from the less privileged, noisier backgrounds.

Our result is a consequence of the nature of the information structure and the process of human capital acquisition we postulate. We assume that both human capital and test scores are affected both by learning at school and by innate ability. However, learning at school has a “comparative advantage” in affecting the school test: an increase in school learning that compensates in the school test exactly for lower ability would not be sufficient to compensate for the reduction in human capital. In other words, given that universities do not observe ability directly but must infer it from the exam results, Spence signalling operates. If a social group has low variance in ability, then a student from that group that obtains a very good exam result is more likely to be considered to have been lucky on the day of the exam than it would be the case for a student from a group with a higher variance in the distribution of ability: the test score is less informative and so in discounted more heavily for groups with low variance in ability distribution. This also implies that effort is less productive for students in low variance groups, and so they will exert lower effort than their peers in groups with higher variance.

The paper contributes to the debate on affirmative action.[3] Affirmative action policies modify admission standards according to observable characteristics of the individuals, such as ethnicity or socio-economic status, and are established in governmental and educational institutions to promote the education and job opportunities of minority or disadvantaged groups. It is a common practice in many U.S. universities (Bowen and Bok, 1998), and in the last decade has been introduced out of the United States: the British government policy promotes access to university by applicants from disadvantaged backgrounds since a decade, while a U.S.-style, race-based affirmative action law has been approved in Brazil in Summer 2012 (BBC, 8^{th} August 2012). In a model of statistical discrimination in the labour market, Coate and Loury (1993) show that, with the introduction of affirmative action policies, a vicious circle may occur between employers, who believe that individuals belonging to a given social group acquire less human capital and pay them less, and individuals belonging to this group who, in turn, have less incentives in acquiring human capital. The equilibrium behaviour that emerges is thus inefficient. On the contrary, De Fraja (2005) supports affirmative action programs, by showing that the optimal provision of education is higher for students coming from socially disadvantaged groups. In the analysis of university admission, Epple *et al.* (2008) and Chan and Eyster (2004) find that, when universities have a preference for diversity, banning affirmative action will significantly decrease the number of minorities in top-tier colleges.

[3]In a recent contribution, Arcidiacono and Lovenheim (2016) survey the empirical literature on the effects of affirmative action programs.

Even though the purpose of affirmative action policies is to sustain minorities or disadvantaged groups, our results support them on efficiency grounds, once ascertained that a discriminated/disadvantaged group has a higher variance in the distribution of ability. Indeed, the test score being necessary for admission in a university would be efficiently lower for a student from a discriminated/disadvantaged group than a student from another social group, as prescribed by affirmative action policies.

Compared to the present analysis, the small literature on admission standards focuses its attention on the asymmetric information problem and does not takes into account differences in observable characteristics that may affect the belief about a student's ability. Pioneering works on admission standards are Costrell (1994) and Betts (1998). Epple et al. (2006) model college admissions as a bargaining game between the college and the potential student under asymmetric information, and analyses the evidence of student "profiling" practices in the college admission processes. Information about student abilities is revealed in the negotiation process. Gary-Bobo and Trannoy (2008) provide a normative analysis of the admission process and emphasise the double-sided asymmetric information problem of a student's ability. They show that this can be solved through a mixed policy of tuition fees and examinations.

Some contributions investigated the effects of the introduction of admission test in a political economy setting. De Fraja (2001) considers a framework where students differ in ability and income, the pay university fees, and there is uncertainty in their future income. Admission tests are implemented together with a uniform subsidy to university attendance financed by a proportional tax on parents' income. He shows the emergence of a participation gap as richer parents are more likely to enroll their child at university, and thus taking the financial risk of uncertain returns to education. In a recent paper, De Donder and Martinez-Mora (2017) extend De Fraja (2001) in a general equilibrium setting, where parents vote on the admission level of universities, then choose whether to invest in private tutoring to help their children to pass the test. They show that a university participation gap emerges between rich and poor students, since richer parents invest more in tutoring.

The present framework is similar to MacLeod and Urquiola (2009, MU hereafter). The information structure is identical, their education technology slightly different, since we allow a school to differentiate their teaching effort among different student types. Moreover, in MU, students enters the job market after school instead of applying for university admission, and are selected by employers. Finally, the analysis of MU is devoted more towards the political economy aspects of school funding and the interaction of different schools (for example state and private, for-profit schools), while we focus on the choice internal to a single school.

The paper is organised as follows. Section 2 contains the model, Section 3 presents the results. Section 4 analyse the case that an anti-discrimination policy is implemented, so that universities cannot use other information than the score in the admission test to evaluate a student's admission. Section 5 concludes the paper.

2 The model

2.1 Students at schools

We study the interaction among students, schools and universities in an education system. There is a large population of students, who are enrolled in school and study there; at the end of the education process, they take an exam, and are subsequently admitted to one of a large number of universities. The focus of our analysis is on the process of admission to university.

The process of acquiring education in schools is modelled along similar lines as in MU's model. In detail, we assume that students enrolled in the schools differ along two dimensions, ability, which is unobservable, and a set of observable characteristics. Ability is measured by a parameter $a \in \mathbb{R}$: that is a student's ability is a unidimensional measure of the set of skills that a student possesses prior to entering school; in general it will be some combination of genetic inheritance (innate ability) and the effect of earlier formative years, whether at home or at kindergarten and primary school. The set of observable variables, which may include sex, social background, race, the type of school attended, the region of residence and so on, is described by a label s, and we assume that s takes one of the possible values in a set $\{1, ..., S\}$. Thus for example, $s = 1$ identifies Caucasian white boys from a working class family living in an run-down area in the North-East who attend a state school, and so on. The label s therefore includes all variables that can be observed and that are not directly related to a person's potential. For the sake of giving it a name, we shall refer to it as "social group".

Following MU, we assume that there is no adverse selection: all participants, students, schools and universities, have the same information regarding the payoff relevant exogenous variables of the model. In particular, the ability of a student, her idiosyncratic a, is known neither to the student herself, nor to the school she attends, nor to the universities she might be applying for.[4] This captures the idea that young individuals, precisely because of their youth and inexperience, may have only vague ideas about their own potential.[5] Although it is not known, information about a person's ability can be gleaned from characteristics which are known to be correlated to ability. Thus we assume that students, schools and universities assess a given student's potential using statistical information regarding that student's social group.[6]

To capture the correlation between social group and ability, we posit that within each group s, innate ability is distributed according to a given and known distribution, which, to facilitate calculations, we take to be normal. The parameters of the ability distributions for each group

[4] This assumption is a simplification, for the sake of tractability, of the more nuanced approach proposed by Gary-Bobo and Trannoy (2008), who assume that the school and the student both observe a separate, imperfectly correlated signal of the student ability.

[5] A slightly richer set-up is used in De Fraja and Landeras (2006), where students know in which quantile of the ability distribution their ability places them. The present model could be extended at the expense of additional complexity.

[6] This is similar to Coate and Loury (1993), and is what insurers do when offering contracts tailored to specific groups.

are common knowledge. Specifically ability in group s is distributed according to

$$\mathcal{N}\left(\bar{a}_s, \frac{1}{\rho_s}\right) \qquad\qquad s = 1, \ldots, S,$$

where $\bar{a}_s$ is the mean ability and $\rho_s = \frac{1}{\sigma_s^2}$ is the precision of a normal distribution with variance σ_s^2. The population comprises a measure g_s of students of type s, with total $\sum_s g_s$ normalised to 1 without loss of generality.

Students attend school in order to be admitted to university and to become better prepared for it. At the end of their school career, students take a test, administered by a central agency.[7] A student's result in this test is assumed to be determined additively by four components: the student's ability, a, the student's effort while at school, e, the quality of the teaching they receive while at school, denoted by b, and a random component, the "luck" on the day of the exam. Formally, at the end of her studies, a student of ability a, who has exerted effort e and who has received schooling of quality b obtains a result t in her test given by

$$t = a + e + \lambda(b) + \varepsilon. \tag{1}$$

In (1), the function λ is increasing ($\lambda'(b) > 0$) and concave ($\lambda''(b) < 0$), reflecting decreasing returns to scale in teaching effort). ε is the realisation of a random variable normally distributed with 0 mean and variance σ_t^2:

$$\varepsilon \backsim \mathcal{N}\left(0, \frac{1}{\rho_t}\right),$$

where $\rho_t = \frac{1}{\sigma_t^2}$ is the precision. Note that the distribution of the random error is independent of the student's social group. We capture differences between social groups through differences in the parameters of the ability distribution. As Proposition 1 shows, the parameters of the distributions that matters enter as a ratio, and so fixing one across groups while allowing the other to vary does not imply a great loss in generality. The assumption of different error distribution for different social groups could also be justified, though it would neither be easier to interpret, nor preferable on grounds of realism.

At the end of the school, students attend university. Admission to university is competitive: universities are free to reject applicants. There is a large number of universities, each choosing a separate set of admission standards. In order to keep the set-up manageable, we assume that universities are fixed in size, that is, they cannot admit more that a given number of student. Universities can select applicants on the basis of observable variables, the results of exams, interviews and tests, and also the student's social group. That is, it is feasible and permissible for a university to admit a student when students from different social groups with better test scores are rejected.

Formally, if university j sets an admission standard τ_s^j, $s = 1, \ldots, S$, a student from group

[7] We use the term "test", but extending its meaning to include other relatively objectively measures, such as a written personal statement or extracurricular activities would be straightfoward.

s can attend this university if and only if she achieves a test score[8]

$$t \geqslant \tau_s^j.$$

The objective of each university is to admit students who are likely to be successful, in the sense of completing their degree with good grades, or some measure of labour market short or long term success. In the short term, universities are assessed on employment success, such as the proportion of graduates who have a job within six months of graduation. In the long term, successful graduates are more likely to donate to alumni funds, or to bestow honour and prestige with recognitions such as knighthoods, honorary degrees and so on. We assume that success is (positively) correlated to a person's human capital when she enters university, $h \in \mathbb{R}$ and to the relevant characteristics of the university she attends, the quality of the teaching, the reputation of the institution, the peer group, the composition of the student body, and the effort of the student while at university. We take all these factors as given and normalise them away. In a richer version of the model, universities could adapt their teaching effort to the students' social group, and test score. This would make their behaviour similar to schools' and effectively split into two separate period the process of human capital formation. Given that the focus of the paper is on the admission process of the universities, this extension would be cumbersome and add little additional insight. By the same token, externalities from within the student body, such as the peer group or diversity, are factored in the individual characteristics of the universities, which affect the students' ranking, but can otherwise be excluded from the model.

Given all its characteristics, to maximise success, each university admission process will need to select students with as high a human capital, h, as possible. h depends on the student's "innate" ability, on her effort at school, and on the quality of the teaching she receives at school. Naturally, these are the same components which affects the result of the test, minus the random "luck on the day" error; also naturally, their relative importance is different for the test score and for the long term human capital. Formally, we let a student human capital be given by:

$$h = a + \eta e + \beta \lambda (b) . \tag{2}$$

Here, $\eta, \beta \in [0, 1]$ are the parameters measuring the relative effectiveness of student effort and teaching in affecting the test and the human capital acquired by the student. The relative importance of ability in the test and in the creation of human capital is normalised to 1. The restriction that η and β be at most 1 implies that, relative to innate characteristics, the student's effort at school and the quality of the teaching she receives are more effective in helping students do well in the test than in helping them to be successful in their post-school career.[9]

[8] This set-up can be easily adapted to the study of the case where universities are prevented from differentiating students on the basis of social background simply by adding the additional constraint that τ_s^j be constant in s.

[9] Note that in MU's set-up, η and β are instead restricted to be exactly 1.

This seems plausible for a variety of reasons. Firstly, different students pursue different careers, and skills acquired at school which help one be, say, a good lawyer might be less useful to be a neurosurgeon: the very fact that we consider a good school one that forms a "well-rounded" young person requires that a balance be struck among several possible sets of skills. Secondly, there a natural depreciation of a person's human capital, due simply to the passing of time: some of what is learnt at 16 is forgotten at 25. Finally, it seems natural to posit that the work of a student and of her teachers' be directed to pass the school exam: this is a result which is immediately measurable in a short space of time.[10] The parameters η and β are familiar ones in the Spence signalling set-up (Spence 1973): $\eta = \beta = 0$ is the extreme case of education as a pure signal, adding nothing to a person's human capital.

2.2 Payoffs

The objective of university j is to admit students that have a "good" level of human capital, to enhance prestige, as mentioned above. It does so by choosing the vector of admission thresholds, $\tau_s^j,\ s = 1, \ldots, S$. Notice that while we have ruled out adverse selection, and therefore, a student's effort while at school is independent of her ability, we allow moral hazard, and in particular, while a student knows her effort level at school, universities do not, and must therefore infer a student's level of human capital from the information at their disposal, her social group and her exam results. University j takes into account the choices of the other universities, of the students and of the schools, and aims to maximise the average human capital of its students; given its fixed size, this is equivalent to maximising the total human capital of its student body.

Consider students next. If a student attends university j, she receives benefit u_j. Her payoff function is the difference between u_j and a function measuring the disutility cost of effort, $\psi(e)$, which has standard properties, $\psi'(e), \psi''(e) > 0$. All students rank universities in the same way.[11] A student takes the behaviour of universities as given. A student who receives an admission offer from more than one university, will simply choose the highest ranked among them. This determines a functional relationship between the test score and the university the student is admitted to. If there are sufficiently numerous universities, a monotonic transformation turns the student's benefit into her human capital.

Finally, we describe the payoff function of the school. While each university takes its decision independently of the others, we model here the schools as a single player, which we refer to as "the schools" or "the school system". This reflect the practice in many jurisdictions, where a central authority, the government, or the school board mandates several aspects of the education policy in its region of competence.

[10] A frequent gripe is that schools teach to the test, forgetting what really matters for education: but this, it could be argued, is a consequence of poorly designed tests, that is, tests that can be passed well only thanks to good exam passing techniques, which are uncorrelated to good measures of a student's human capital.

[11] Relaxing this assumption creates a richly complex model, but one whose questions of interest are tangential to the present paper.

The payoff of the schools and the constraints that the schools must obey are in general affected by policy and the preferences of the individuals running the schools. In order to keep a general set-up, we posit that the schools choose the vector[12] of teaching effort levels $\mathbf{b} = \{b_1, \ldots, b_S\}$ and receive a benefit given by

$$\sum_s g_s u_s(\mathbf{b}) + \mu \sum_s g_s m_s \bar{t}_s, \tag{3}$$

and are subject to the following constraints:

$$B + \sum_s g_s m_s \bar{t}_s \geqslant c\left(\sum_s g_s b_s\right), \tag{4}$$

$$b_s \geqslant \bar{b}. \tag{5}$$

In (3), $u_s(\mathbf{b})$ is the (average) university attended by students of social group s, $s = 1, \ldots, S$, so the first component of the school payoff (3) is the average success of the students measured by the university they attend. The second term captures the benefit of successful exam results. $\bar{t}_s$, $s = 1, ..., S$, is the average test score obtained by students in group s.

Constraint (4) is the budget constraint: on the RHS of (4), there is the school cost, given by an increasing and convex function of the total amount of teaching:

$$c\left(\sum_s g_s b_s\right), \tag{6}$$

with $c'(\cdot), c''(\cdot) > 0$. Revenues are on the LHS of (4): a fixed component, B, and a performance related one, the second addendum. This captures the assumption that the school receives a sum $m_s t_s$ when a student from social group s obtains a test score t_s. As implied in (3), this additional budget of the school also enters its payoff function, possibly because the school likes to have a large budget, or because it derives (non-monetary) utility when the exam results of its students are good. The parameter μ measures the relative weight of university success (measured by the "average university" this school's pupils are admitted to), and short term exam results. The payoff function in (3) is quite general,[13] and captures various special cases: thus $\mu = 0$ implies that the schools have no direct preference for results: results only affect schools via their effect on the budget. Letting all the m_s's go to 0 captures the case of a fixed budget (any preference for results would need to be captured by μ tending to ∞, with the term $\mu \sum_s g_s m_s$ instead tending to a positive value).

[12] We note here a difference with the set-up proposed in MU. The term $\lambda(b)$ differs from the corresponding term in expression in MU (p 6), β, which captures individual characteristics of the school, the "school value added", and therefore it has the same value for all students in a given school. In our paper, it can differ from student to student in the same school, because it reflects the choices made by the school regarding the allocation of its resources.

[13] Adding further generality by assuming different relative weights in the objective function and in the budget formula seems an unnecessary and unrewarding complication.

Constraint (5) binds b_s from below: each child must receive a minimum acceptable teaching quality. Once this standard is met, schools can use their discretion to allocate their remaining resources to some students only. They can, for example, provide additional lessons, which require longer opening hours for buildings, and additional staff, or overtime pay to existing staff; or they can hire more highly qualified personnel in certain subjects, or send some of their staff to non-mandatory training, or purchase higher quality equipment and so on. By devoting more resources to subjects studied by certain social groups, or target them more directly, for example, with help for out-of pocket expenses or the costs of additional lessons, schools can tailor the quality the teaching provided to the observable characteristics of the students.[14]

3 Results

To sum up, in the game described above, universities commit to their admission policies; knowing these policies (which are anyway time consistent), students and schools choose their effort. As we explain below, it is immaterial whether sequentially or simultaneously. Students and schools also understand the way other students and schools make their choices, and so form correct belief about the effort exerted by the other students in the school and the teaching effort devoted by the school to each of them. Information about these variables is important to a given student, because universities take them into account, when assessing the effort a student (is believed to have) exerted at school, and the teaching effort the school devoted to her.

3.1 Students' effort

Given the payoff function of a university, its optimisation problem is straightforward. In equilibrium, it knows the distribution of human capital in each social group, and therefore it knows the functional relationship between the test score obtained by a student in a given social group and her *expected* human capital; they also know the students' ranking of universities and hence which university each student will choose among those which have made her an offer. It cannot be optimal for a university to admit a student i_0 when there is a student i_1 who is not admitted, and who has, in expectation, higher human capital than student i_0. It follows that each university simply makes an offer to a student if she is believed to have ability above a certain threshold. The belief is formed taking the expectation conditional on all available information, namely the student's test score and her social group. The threshold is calculated to ensure that the correct number of places is filled by candidates whose h is above the threshold but below the threshold set by the university immediately above it in the students' ranking.

Figure 1 illustrates. It depicts the densities of universities' belief about human capital in

[14] In practice of course, some of these acivities are provided specifically for children of certain abilities, not necessarily low achievers: while some teachers are inclined to spend more attention on weaker students, others might relish the opportunity to help bright students to excel. It is a limitation of our model is that it cannot allow the school system to skew expenditure towards children of specific abilities.

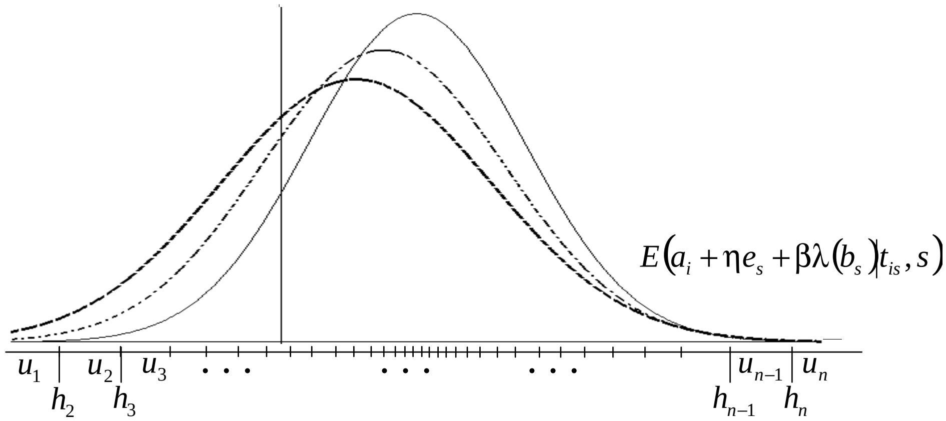

Figure 1: Distribution of human capital in three social groups.

three groups of students. Universities know these densities, since they have correct beliefs about the ability, effort and teaching quality, and "day-of-the-exam" error in these groups of students. Students, from all three groups, who are believed to have human capital exceeding h_n receive an offer from every university, but, given their preferences, enrol in university u_n. University u_n selects its threshold so that the number of students with test score at least h_n is equal to the number of places it has on offer. Next students whose human capital is believed to be between h_{n-1} and h_n receive the offer of a place from all universities except u_n, and so end up in university u_{n-1}, their preferred one among those which make them an offer. And so on for all students, down to those whose human capital is believed to be below h_2, who receive an offer only from university 1.[15] In the limit, which is the case we consider, the ability intervals in Figure 1 are infinitely small. In equilibrium, universities are fully stratified, with the university who is considered the most preferred by students admitting the highest human capital students (in measure equal to the number of places it has), the second the next interval in the distribution of expected human capital, and so on. Knowing this, students choose their effort so as to maximise their "expected university". This, for student i from social group s, who obtained t_{is} in the test, is given by:

$$E\left(a_i + \eta e_s + \beta\lambda\left(b_s\right)|t_{is}, s\right). \tag{7}$$

where e_s and b_s are equilibrium levels of effort exerted by students in social group s and the teaching effort devoted by schools to those students.

From Bayes rule (De Groot 1970, p 167), we can expand (7) to derive the following expression for a student's human capital and hence the university she can attend:

$$u_{is} = \frac{\rho_s}{\rho_s + \rho_t}\left(\bar{a}_s + \eta e_s + \beta\lambda\left(b_s\right)\right) + \frac{\rho_t}{\rho_s + \rho_t} t_{is}. \tag{8}$$

[15] The case where not all students attend university, but receive a labour market payoff which is increasing with their human capital can be easily incorporated in this framework.

If that student exerts effort e_i, we can use (1) to derive her expected test score, and so obtain her university as a function of her chosen effort level:

$$u_{is}(e_i) = \frac{\rho_s}{\rho_s + \rho_t}(\bar{a}_s + \eta e_s + \beta\lambda(b_s)) + \frac{\rho_t}{\rho_s + \rho_t}(\bar{a}_s + e_i + \lambda(b_s)). \tag{9}$$

It is now straightforward to determine the student's optimal choice of effort.

Proposition 1 *A student from social group s exerts effort e_s given by:*

$$\psi'(e_s) = \frac{\rho_t}{\rho_s + \rho_t}. \tag{10}$$

Proof. Given universities' equilibrium belief regarding the effort exerted by a student in her social group, a student chooses her own effort e_i to maximise the difference between (9) and the cost of effort $\psi(e_i)$:

$$\frac{\rho_t}{\rho_s + \rho_t} = \psi'(e_i).$$

In equilibrium, $e_i = e_s$, which gives (10). The second order condition, is satisfied as $-\psi''(e_s) < 0$. ■

Thus the effort exerted by a student depends on the relative precision of the random error in the test and of the distribution of ability in her social group. To the extent that this differs across groups, students from different social groups exert different levels of effort. This is natural, once it is clear the dual role of effort: it both increases human capital directly, and, as Spence (1973) pointed out, it affects the signal and hence a university's belief about "innate" ability, which is a component of human capital. But the benefit of trying to alter the signal depends on the accuracy of the signal: in the extreme case where all students in a given social group s are of similar ability (ρ_s is very high), universities who see a student with a very high score will assume that she has been lucky in her test, since they assume that she has exerted the "average" effort for group s, and so universities will not be strongly influenced by the test in assessing the student's ability. Consequently, students in groups s derive little benefit from their effort and therefore have little incentive to exert it. Conversely, if the test is instead a very accurate measure of ability (high ρ_t), then affecting it will be very important. With high ρ_t effort is a very effective way of improving one's post-school prospects, and the student has a strong incentive to exert effort.

A university assessing a student with a unexpectedly (given her social group s) good test score will attribute the good result to a combination of high ability and good luck. If it is very unlikely that the student has very high ability (because her social group has low variance, and hence very small tails) then her good test score will be attributed mostly to good luck, with little effect of the assessment of the student. Consequently this student will have low incentive to exert effort, as trying to "outperform" her peers will not influence much the university. The reason is similar to the intuition underlying Dewatripont-Jewitt-Tirole (1999a, 1999b)'s explanation of the lower effort level exerted by managers who are given multiple tasks: a good performance is more likely to be attributed to luck than to ability.

It is interesting to compare the effort level derived in (10) with the efficient level of effort. This is the level that would be exerted without the distortion caused by unobservability of ability, and is straightforward to calculate. A student whose ability is perfectly observable to the universities, would choose e_i to maximise $a_i + \eta e_i + \beta\lambda(b_s) - \psi(e_i)$ and so choose e^* satisfying:

$$\psi'(e^*) = \eta, \tag{11}$$

which, unlike the value derived in (10) when ability is unobserved, is independent of the social groups and of the parameters of the distribution of the test error. Note also that effort can be inefficiently *high*: this happens if $\eta < \frac{\rho_t}{\rho_s+\rho_t}$, as in Spence's (1973) signaling model, where $\eta = 0$.

Students' effort does *not* depend on teaching effort. In general, one would expect students to respond to their school's teaching: for example, De Fraja and Landeras (2006) study a game in choice of effort between students and schools. They show that the sign of the effect of school effort on students' effort is ambiguous, as students may respond to increased effort on the schools' part either by reducing or increasing their own effort; that is, effort choices can be strategic substitutes or complements. Our conclusion in Proposition 1 where school effort is absent, therefore, sits right in the middle of this ambiguity. From an analytical viewpoint, it has the additional advantage that it makes it irrelevant whether the school and the students chooses simultaneously or sequentially. Thus school effort can be interpreted either as the quality of classroom interaction, which would be set simultaneously with students' choices, or as investment, in, say, computers, classroom equipment teachers' training, which is set in advance and is fixed by the time students choose their own effort.

3.2 Effort and social group

Proposition 1 shows that students from different social groups may exert different effort levels, and consequently may obtain different exam results even when they have the same ability and are subject to the same "day-of-the-exam" error. So far, we have treated social groups as if they were labels. In practice, of course, social groups exhibit systematic differences, which can form the basis of policy. Groups, in particular, have different ability distributions; as a convention, we index the social groups ordering them in order of average ability, so that $\bar{a}_s > \bar{a}_{s-1}$ for $s = 2, ..., S$.

A consistent empirical regularity across the world is a positive correlation between ability and social background.[16] Thus, in the rest of the paper, a higher s, associated with higher

[16] The association between social class and children's cognitive abilities is empirically very well documented. Students with a disadvantaged social background enter primary school behind their more advantaged peers (Marton *et al.*, 1997; Mayer 1997; Lee and Burkam 2002), and these initial cognitive differences increase as children progress through school (Phillips *et al.*, 1998; Reardon 2003; Downey *et al.*, 2004). An early crop of studies (Binet and Simon, 1916; Burt, 1922; Davis, 1948; Eells *et al.*, 1951; MacArthur and Elley, 1963; Cropley,

average ability, denotes a "better" social background.[17]

As (10) shows, it is the variance, not the mean, of the distribution that affects effort. While the body of empirical evidence is less ample than for the mean ability, what empirical evidence there is does suggest that social groups with higher ability also have less dispersion in ability. For example, in an early UK longitudinal studies, the measure of innate ability used by Galindo-Rueda and Vignoles (2003) and De Fraja *et al.* (2010) among others, shows that the ability distribution of children from higher socio-economic groups has higher mean and lower variance than that of children from less favoured households. In a more indirect test, Hauser (2002) regresses mean against standard deviation in several occupational groups, and find a weak negative relationship.[18]

Formally we posit the following in the rest of the paper.

Assumption 1 $\bar{a}_{s_1} > \bar{a}_{s_2}$ *implies* $\rho_{s_1} > \rho_{s_2}$.

With abuse of notation we will occasionally treat s as a continuous variable, and denote by $\bar{a}'(s)$ and $\rho'(s)$ the differences $\bar{a}_s - \bar{a}_{s-1}$ and $\rho_s - \rho_{s-1}$, $s = 2, ..., S$. Our labeling of groups in increasing order of ability implies $\bar{a}'(s) > 0$; this and Assumption 1 imply $\rho'(s) > 0$: groups with higher ability also have lower variance.

The following is an immediate consequence of Proposition 1.

Corollary 1 *If Assumption 1 holds, then*

$$\frac{de_s^*}{ds} = -\frac{\rho_t \rho'(s)}{\psi''(e_s^*)(\rho(s) + \rho_t)^2} < 0. \tag{12}$$

This follows from total differentiation of (10) and says that students from a group with higher average ability will exert less effort. The intuition can be again traced back to Dewatripont-Jewitt-Tirole (1999a, 1999b): a good test score is more likely to be due to luck for a student whose ability is better known. A university assessing a student with a unexpectedly good test score (given her social group s) will attribute the good result to a combination of high ability and good luck. If it is very unlikely that the student has very high ability (because her social group has low variance, and hence very small tails) then her good test score will be attribute mostly to good luck, with little effect of the assessment of the student. Consequently this student will have low incentive to exert effort, as trying to "outperform" her peers will not influence much the university.

1964; Kenneth and Cropley, 1970) investigated the relationships between results of intelligence tests and social class and showed a positive correlation between intelligence (measured by IQ) and socio-economic status.

[17] Of course, for some characteristics, such as sex, the groups are not necessarily ordered, though they would be for example, in societies where earlier schooling is differentiated for boys and for girls.

[18] Qualitatively similar results in an early childhood US longitudinal sample are found by Morgan *et al.* (2009), in the analysis of learning difficulties of children of 24 months of age differing in socio-economic status: children from lower socio-economic status households are about twice as likely as those from high socio-economic status households to show learning problems.

3.3 School's effort

To determine the equilibrium school effort, and hence determine the last component of a student's human capital, we need to solve the school's optimisation problem (3). This we do in the next proposition; with the same abuse of notation as before, we let $m'(s) = m_s - m_{s-1}$.

Proposition 2 *At the solution of (3):*

$$\frac{db_s}{ds} = \frac{(\lambda'(b_s))^2}{\lambda''(b_s)\, c'(\cdot)} \frac{\left(\frac{c'(\cdot)}{\lambda'(b_s)} - m_s\right) \frac{(1-\beta)\rho_t}{(\rho_s+\rho_t)^2} \rho'(s) - \left(\frac{\rho_t+\beta\rho_s}{\rho_s+\rho_t} + \mu \frac{c'(\cdot)}{\lambda'(b_s)}\right) m'(s)}{\frac{\rho_t+\beta\rho_s}{\rho_s+\rho_t} + \mu m_s}. \tag{13}$$

Proof. See the appendix. ■

Expression (13) in Proposition 2 is complex and not immediately interpretable. However, it can be used to infer the effect on the equilibrium values of the school effort of changes in the exogenous parameters, in some special relevant cases. As a first such case, suppose that schools are simply given a performance related bonus based on the *average* exam result: the performance related part of the budget is independent of the social class of the student. In this case, m_s is constant, and putting $m'(s) = 0$ in (13), it becomes:

$$\frac{db_s}{ds} = \frac{\lambda'(b_s)}{\lambda''(b_s)} \frac{\frac{(1-\beta)\rho_t}{(\rho_s+\rho_t)} \rho'(s)}{\left(\frac{\rho_t+\beta\rho_s}{m} + \mu\right)} \left(\frac{\lambda'(b_s)}{c'(\cdot)} - \frac{1}{m}\right). \tag{14}$$

To interpret this, note that $\lambda''(b_s) < 0$, that $\beta < 1$ and that $\rho'(s) > 0$, by Assumption 1. So the sign of $\frac{db_s}{ds}$ is the opposite of the sign of

$$\lambda'(b_s) - \frac{c'(\cdot)}{m} \tag{15}$$

When b_s is low, then $\lambda'(b_s)$ is high, and (15) is likely positive, making $\frac{db_s}{ds}$ positive. Thus when m is small, $b'_s(s)$ is more likely to be increasing with s, which means that students from groups with higher average ability receive more teaching. *Vice versa* for higher m.

A different simplification is obtained if $\beta = 1$. We imposed $\beta < 1$, since it seems plausible that school quality is relatively more important than innate ability for test scores than for long term human capital. In the extreme case when teaching is instead as important as ability for human capital, (13) simplifies to

$$\frac{db_s}{ds} = -\frac{(\lambda'(b_s))^2}{\lambda''(b_s)\, c'(\cdot)} \frac{\left(1 + \mu \frac{c'(\cdot)}{\lambda'(b_s)}\right) m'(s)}{1 + \mu m_s} \tag{16}$$

which has the same sign of $m'(s)$.

Consider next the case in which ability is known to all agents. Since universities are aware of a student's ability, now the distribution of ability in each social group is not relevant in the admission process. On the other hand, knowing a student's social group is important for universities if the school differentiates teaching effort along social groups. Moreover, given the group of students of type (a, s) and observing their average result in the test, universities can also ignore the luck in the test ε, since they know that the level of human capital of all the students of type (a, s) is the same. Given the behaviour of universities, all the students exert the same effort. Indeed the expected university of a student does not depend anymore on the distribution of ability within a social group, nor on the precision of the test. In other words, a student's effort is independent of the social groups and of the parameters of the distribution of the test error. Also, now teaching effort can change according to both ability and social group, i.e., $b_{as} = b(a, s)$. However, given the school's objective function, which is sensitive to a student's social group but not to a student's ability, all the students belonging to a certain social group would receive the same teaching effort.[19] In particular, the intensity of teaching effort will depend on whether the performance related bonus m_s is higher for students belonging to a certain social group s. Note that if b^*_{as} was the same for every s, then universities would only look at results in the test. Therefore the signaling role of the social group is not related on the distribution of ability but on the differences in teaching effort according to it. Indeed, even if universities cannot directly observe teaching effort, they can infer it by observing the (average) test scores of students belonging to different social groups and having the same ability: if the average test score is different, a university is aware of the differences in teaching effort along social groups.

Thus a student i's expected university is

$$u_{is}(e_i) = t - (1-\eta) e^* - (1-\beta) \lambda (b^*_{as}),$$

where e^* is such that equation (11) holds. If a student obtains a result t, her ability is:

$$a = t - e^* - \lambda (b^*_{as}),$$

and so will have human capital

$$\begin{aligned} h &= a + \eta e^* + \beta\lambda (b^*_{as}) \\ &= t - (1-\eta) e^* - (1-\beta) \lambda (b^*_{as}). \end{aligned}$$

Hence the school's problem is

$$\max_{\{b_{1a},\ldots,b_{Sa}\}} \sum_s g_s (\bar{a} + \eta e^* + \beta\lambda (b_{as})) + \mu \sum_s g_s m_s (\bar{a} + e^* + \lambda (b_{as})), \tag{17}$$

[19] Nonetheless, even assuming for the sake of the argument that the school's objective function is such that teaching effort may change according to a student's ability also, this situation is very unlikely to occur in the real world, as the cost of differentiating teaching effort for each student becomes extremely high.

$$\text{s.t.: } B+\sum_s g_s m_s\left(\bar{a}+e^*+\lambda\left(b_{as}\right)\right)\geqslant c\left(\sum_s g_s b_{as}\right),$$
$$b_{as}\geqslant\bar{b}.$$

The first order conditions of (17) are:

$$\lambda'\left(b_{as}\right)\left(\beta+m_s\left(\mu+\nu\right)\right)=\nu c'\left(.\right), \qquad s=1,...,S,\ a\in\mathbb{R} \tag{18}$$

Total differentiation of the LHS of the above w.r.t b_{as} and s gives:

$$\lambda''\left(b_{as}\right)\left(\beta+m_s\left(\mu+\nu\right)\right)db_{as}+\lambda'\left(b_{as}\right)\left(\mu+\nu\right)m_s'ds=0,$$
$$\frac{db_{as}}{ds}=-\frac{\lambda'\left(b_{as}\right)}{\lambda''\left(b_{as}\right)}\frac{\left(\mu+\nu\right)m_s'}{\beta+m_s\left(\mu+\nu\right)}. \tag{19}$$

Teaching effort is higher the higher the performance budget for a social group. If m_s is higher for the test scores of students from a specific social group, then more teaching effort will be exerted towards those students. Thus the sign of (19) will depend on whether the school has more incentives in exerting teaching towards students from a low or high social group, i.e.,

$$\frac{db_{as}}{ds}\propto m_s'.$$

Notice that, given a performance related bonus based on the *average* exam result ($m_s'=0$), then b_{as}^* would be the same for all students, leading to $\frac{db_{as}}{ds}=0$. Conversely, total differentiation of the LHS of (18) with respect to b_{as} and a gives:

$$\lambda''\left(b_{as}\right)\left(\beta+m_s\left(\mu+\nu\right)\right)db_{as}=0,$$
$$\frac{db_{as}}{da}=0.$$

3.4 University selection

In this section we consider the behaviour of universities. The main result in the paper is in the following proposition.

Proposition 3 *Let Assumption 1 holds. Let $u_s(t)$ be the university attended by students from social group s who leave school with a test score $t\in\mathbb{R}$. Then there exists t_s^* such that $u_s(t)\gtreqless u_{s-1}(t)$ according to $t\lesseqgtr t_s^*$.*

Proof. This follows from (8). The relationship between the result obtained in the test, and the university attended is a straight line with slope $\frac{\rho_t}{\rho_s+\rho_t}$. Since, by Assumption 1, $\rho_{s-1}<\rho_s$, the slope is lower than the slope for a lower social group $s-1$. ■

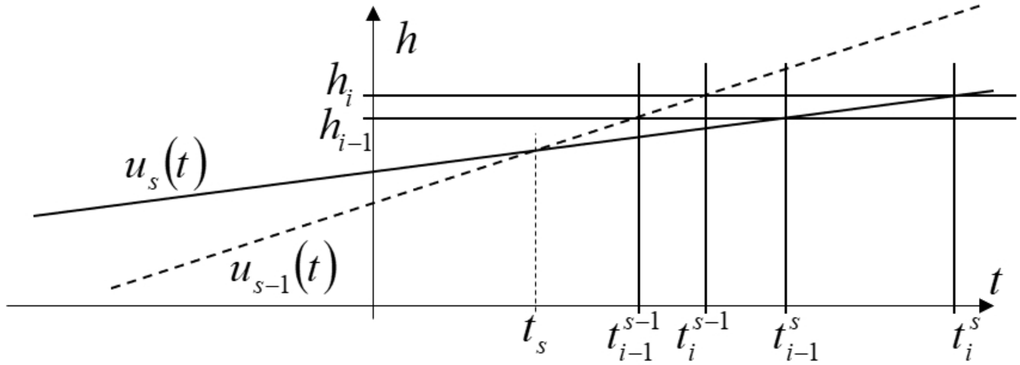

Figure 2: Relationship between human capital and test result.

Figure 2 illustrates the situation. It compares students from two social groups, s and $s-1$. The dashed line is the level of human capital that a student from social groups $s-1$ whose test score is t is believed to have, measured on the horizontal axis, and hence the highest ranked university that makes her an offer. The solid increasing line is the corresponding relationship for student from social group s. The university whose students have human capital between h_{i-1} and h_i, will make offers to any student from social group $s-1$ whose exam result is at least t_{i-1}^{s-1}, and from social group s whose test score is at least t_{i-1}^{s}. Therefore students who study at this university will have passed their school exam with a score in $\left[t_{i-1}^{s-1}, t_i^{s-1}\right]$ if they are from social group $s-1$ and with a higher score with a result in $\left[t_{i-1}^{s}, t_i^{s}\right]$ if they are from social group s. The result is reversed for students who achieve a test score below t_s^*, that is, for less competitive universities.

The diagram also illustrates clearly the reason why the students in the group with higher variance exert more effort: a given improvement in the exam result translates into a greater improvement in the most preferred university that makes them an offer than for students from the group with lower variance, and so their incentive to exert effort is stronger.

When ability is not observable, the effect of ability on the neat relationship is mixed with the signalling effect of the test, and with the different effort exerted by students from different social groups and the different quality of the teaching they receive. Thus, for example, if $\beta < 1$ and $\rho'(s) > 0$, then as we saw, students from higher social groups exert less effort, and their teachers devote less effort to them. However, they are on average abler, and therefore their average salary will be higher or lower depending on the balance of these effects. To see this, differentiate the expected human capital of a student from social group s and rearrange using (12) and (13):

$$\begin{aligned} u_{is}(e_i) &= \frac{\rho_s}{\rho_s+\rho_t}\left(\bar{a}_s + \eta e_s + \beta\lambda(b_s)\right) + \frac{\rho_t}{\rho_s+\rho_t}\left(\bar{a}_s + e_i + \lambda(b_s)\right) \qquad (20) \\ &= \bar{a}_s + \frac{\eta\rho_s+\rho_t}{\rho_s+\rho_t}e_s + \frac{\beta\rho_s+\rho_t}{\rho_s+\rho_t}\lambda(b_s). \end{aligned}$$

Corollary 2 *Differentiation of (20) with respect to s yields*

$$u'(s) = \bar{a}'_s - \left[\rho'_s \rho_t \left(e_s (1-\eta) + \lambda (b_s(s)) (1-\beta)\right)\right] (\rho_s + \rho_t)^{-2} - \frac{(\eta \rho_s + \rho_t) \rho_t \rho'(s)}{\psi''(e^*_s)(\rho_s + \rho_t)^3} \tag{21}$$
$$- \frac{\beta \rho_s + \rho_t}{\rho_s + \rho_t} \frac{(\lambda'(b_s))^3}{\lambda''(b_s) c'(\cdot)} \frac{\left(\frac{c'(\cdot)}{\lambda'(b_s)} - m_s\right) \frac{(\beta-1)\rho_t}{(\rho_s+\rho_t)^2} \rho'(s) + \left(\frac{\rho_t + \beta \rho_s}{\rho_s + \rho_t} + \mu \frac{c'(\cdot)}{\lambda'(b_s)}\right) m'(s)}{1 + \frac{\rho_s + \rho_t}{\rho_t + \beta \rho_s} \mu m_s}$$

Proof. See the appendix. ■

The difference in average university between two social groups is the difference in their innate ability, reduced or increased by a factor of exogenous parameters, the terms in the square brackets. The complexity in (21) reflects the various factors influencing the admission decision by universities to students with different characteristics. Clearly, the link between social group and expected university is potentially ambiguous. Students from a group with higher average ability may not gain admission to a given university, because of a combination of lower learning effort on their part and lower teaching effort on their school's part. As a benchmark case, suppose that $\eta = \beta = 1$. Then, (21) reduces to

$$\bar{a}'_s - \left(\frac{\rho_t \rho'(s)}{\psi''(e^*_s)(\rho(s) + \rho_t)^2} + \frac{(\lambda'(b_s))^3}{\lambda''(b_s) c'(\cdot)} \frac{\left(1 + \mu \frac{c'(\cdot)}{\lambda'(b_s)}\right) m'(s)}{1 + \mu m_s} \right).$$

This is negative if the variance of the higher ability group is sufficiently smaller ($\rho'(s)$ high), or if the disutility cost of effort is sufficiently close to being linear ($\psi''(\cdot)$ small), if the test is very accurate (ρ_t large) and if the relative weight of future career success μ or the shadow price of the budget constraint ν is sufficiently high.

We conclude the section by considering the case with complete information. Now universities observe each student's ability directly, so thata student's expected university is:

$$u_i = a_i - (1-\eta) e_i - (1-\beta) \lambda (b_{as}) + t_{is} (\lambda (b_{as})). \tag{22}$$

Teaching effort may change along different social groups, so that an increase in teaching effort lowers the intercept of (22) and increases the test score, i.e., $-(1-\beta) \lambda'(b_{as}) db_{as} < 0$ and $t'_{is}(\lambda'(b_{as})) db_{as} > 0$. Thus the relationship between test score and university admission strictly depends on the design of incentives along different social groups. A student belonging to a social group that receives more teaching effort needs a lower grade to be admitted into a university with a certain quality, compared to a student coming from a social group that receives less teaching effort. In other words, if the school has an incentive in providing more teaching effort towards a specific social group, then the expected university of a student belonging to that group would be of higher quality. If conversely the school's performance related bonus is based on the *average* exam result ($m'_s = 0$), so that the school provides the same teaching effort to all social group, then a student's social group plays no role in determining the relationship between test scores and university admission.

4 Anti-discrimination policy

In this section we assume that universities are prevented from setting different admission standards across social groups. In this case, the test score is the only signal of ability that can be considered by universities. As a consequence, the expected test score will represent the level of university in which a student expects be admitted, i.e.:

$$u_{is} = \bar{a}_s + e_i + \lambda (b_s) . \tag{23}$$

Furthermore, the expression for a student's expected human capital (8) does not represent anymore the expected university:

$$h_{is} = \frac{\rho_s}{\rho_s + \rho_t} (\bar{a}_s + \eta e_s + \beta \lambda (b_s)) + \frac{\rho_t}{\rho_s + \rho_t} t_{is}, \tag{24}$$

where $t_{is} = u_{is}$. By substituting (23) into the students' and schools' optimisation problem, we can obtain the result for the case with no discrimination across social groups.

Begin by students. A student i's problem becomes

$$\max_{e_i} \quad \bar{a}_s + e_i + \lambda (b_s) - \psi (e_s) ,$$

and her optimal choice of effort is given by

$$1 = \psi' (e_s) .$$

Notice that now all students exert the same level of effort. Also, all of them exert higher effort than in the baseline case in which universities can discriminate across students, since the marginal benefit of education is now higher:

$$\frac{\rho_t}{\rho_s + \rho_t} < 1.$$

Consider next schools. Plugging (23) into the school's optimisation problem (3) yields

$$\max_{\mathbf{b}} \sum_s g_s (\bar{a}_s + e_i + \lambda (b_s)) + \mu \sum_s g_s m_s \bar{t}_s, \tag{25}$$

subject to:

$$B + \sum_s g_s m_s \bar{t}_s \geqslant c\left(\sum_s g_s b_s\right),$$
$$b_s \geqslant \bar{b}.$$

Proposition 2 becomes:

Proposition 4 *At the solution of (25):*

$$\frac{db_s}{ds} = -\frac{(\lambda'(b_s))^2}{\lambda''(b_s)\, c'(\cdot)} \frac{\left(1 + \frac{c'(\cdot)}{\lambda'(b_s)}\mu\right) m'(s)}{1 + \mu m_s}. \tag{26}$$

Proof. See the appendix. ■

From Proposition 4 it is possible to examine the differences between (26) and (13). When universities cannot take into account a student's social group, then the differences in the distribution of ability among social groups are not relevant in a school's optimisation problem. Indeed this would not have any effect in a student's admission possibilities. What does still affect a school teaching effort is the differences in performance-related budget across social groups, $m'(s)$. If the performance-related budget is the same for all social groups, then a school has no incentive in modifying its teaching effort according to students from different social backgrounds. Conversely, the school's behaviour will depend on whether the performance-related budget is stronger for some social groups and weaker for other. In other words, the difference in teaching effort according to social groups will be based on the sign of $m'(s)$.

We turn finally on the universities behaviour. University j sets an admission standard $\tau_s^j = \tau^j$, for all $s = 1, \ldots, S$. Therefore a student can attend this university if and only if she achieves a test score $t \geqslant \tau^j$. In equilibrium, all students with grade $t = \tau^j$ will attend university j, all students with grade $t = \tau^k$, $\tau^k = \tau^j + \varepsilon$ will attend university k, and so on.

In the baseline model, all students attending the same university had the same expected human capital, but different results in the test, according to Proposition 3. In this case, given the fact that a university cannot discriminate according to social groups, the results in the admission need to be the same, then all student with same grade will attend the same university level, regardless of their social group. However, their human capital will be different.

Proposition 5 *Let Assumption 1 holds. Let $h_s(t)$ be the expected human capital of a student from social group s who leaves the school with a test score $t \in \mathbb{R}$ and attends university $u = t$. Then there exists u^* such that $h_s(u) \gtreqless h_{s-1}(u)$ according to $t \lesseqgtr u^*$.*

Figure 3 represents the relationship between universities and human capital. Unlike the baseline case depicted in Figure 1, in which all the students attending a university were accepted for admission according to a different result in the test and thus have the same expected human

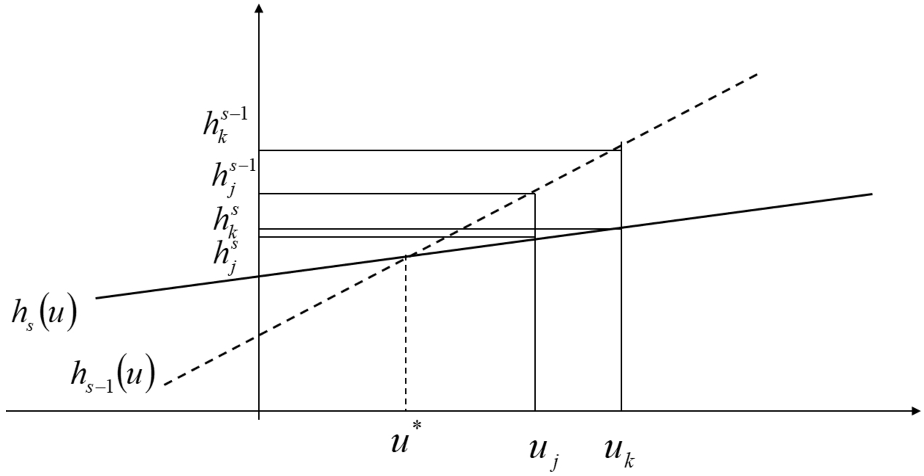

Figure 3: Relationship between human capital and universities

capital, here each university of a certain quality requires the same result in the test, and hence it is attended by students differing in human capital. For instance, university u_k is attended by students from social group $s-1$ with human capital h_k^{s-1} and for students from social group s with human capital h_k^s, where $h_k^{s-1} > h_k^s$.

Notice that, unlike the baseline model, the result in which students coming from a lower social group have higher human capital than students from a higher group (i.e., the area at the right of u^*) is stronger the higher the quality of universities. Moreover, as the university quality increases, so does the difference in human capital of students from different social groups.

5 Concluding Remarks

In this paper we have examined the how possibly discriminative characteristics, such as sex, race or social background, may influence university admission. To do so, we have studied the interplay between students and schools, in terms of learning and teaching effort, to increase students' opportunities of admission. The results show that more competitive universities tend to require lower grades in the admission test to students coming from a disadvantaged group. The opposite applies with less competitive universities. If an anti-discrimination policy is implemented, all students with same grade will attend the same university level, but they will exhibit different levels of human capital based on their social group.

The analysis carried out can be adapted for the analysis of job recruitment rather than university admission, in explaining the role of observable characteristics which are not relevant for the job, such as race, sex, or the family social background in the employers' selection. Field and experimental evidence collected since Goldin and Rouse's (2000) seminal work does indeed strongly suggest that if employers are prevented from learning characteristics deemed irrelevant to the performance of the job, their hiring decisions are altered, often substantially.[20] Indeed recruitment services are proliferating offering employers protection against discrimination by filtering out all "non-relevant" information provided by job applicants, before passing the CV to employers.[21] Our results can be translated to a scenario in which students after school enter the labour market, and the human capital obtained represents the student's salary. Accordingly, students who achieve a high test scores receive a higher salary if they are from the social group with more variance in ability and exert more effort, as for them the incentive to do so is stronger, since a given improvement in the exam result translates into a greater improvement in salary than for students from the group with lower variance.

[20] See also Edin and Lagerström (2006) and Aslund and Skans (2007) for further evidence.

[21] For example, www.2apply.co.uk.

References

[1] Antonji, J.G., and Pierret, C.R. 2001. Employer Learning and Statistical Discrimination. *Quarterly Journal of Economics* **116**: 313-350.

[2] Arcidiacono, P. and Lovenheim, M. 2016. Affirmative Action and the Quality-Fit Trade-Off. *Journal of Economic Literature* **54**: 3-51.

[3] Arrow, K.J. 1973. The Theory of Statistical Discrimination. in Ashenfelter and Rees (eds), *Discrimination in Labor Markets*.

[4] Åslund, O. and Skans, O.N. 2007 Do Anonymous Job Application Procedures Level the Playing Field? Institute for Evaluation of Labour Market and Education Policy.

[5] Backes-Gellner, U. and Veen, S. 2008 The Consequences of Central Examinations on Educational Quality Standards and Labour Market Outcomes. *Oxford Review of Education* **34**: 569-588.

[6] BBC News. 2012. Brazil approves affirmative action law for universities. http://www.bbc.co.uk/news/world-latin-america-19188610

[7] Betts, J.R. 1998 The Impact of Educational Standards on the Level and Distribution of Earnings. *American Economic Review* **88**: 266-275.

[8] Binet, A. and Simon, T. 1916. *The Development of Intelligence in Children*. Williams & Wilkins, Baltimore.

[9] Bowen, W.G. and Bok, D. 1998. *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*. Princeton University Press, Princeton.

[10] Burt, C.L. 1922. *Mental and Scholastic Tests*. P.S. King and son. London.

[11] Chan, J. and EYSTER, E. 2004. Does Banning Affirmative Action Harm College Quality? *American Economic Review* **94**: 858-72.

[12] Chaudhuri, S. and Sethi, R. 2008. Statistical Discrimination with Peer Effects: Can Integration Eliminate Negative Stereotypes? *Review of Economic Studies* **75**: 579-596.

[13] Coate, S. and Loury, G.C. 1993. Will Affirmative-Action Policies Eliminate Negative Stereotypes? *American Economic Review* **83**: 1220-1240.

[14] Costrell, R.M. 1994 A Simple Model of Educational Standards. *American Economic Review* **84**: 956-971.

[15] Currie, J. and Thomas, D. 1999 Early Test Scores, Socioeconomic Status and Future Outcomes. NBER working paper 6943.

[16] Davis, A.1948. *Social Class Influences upon Learning*. Harvard University Press. Cambridge.

[17] De Donder, P. and Martinez-Mora, F. 2017. The political economy of higher education admission standards and participation gap. *Journal of Public Economics* **154**: 1-9.

[18] De Fraja, G. 2005. Reverse Discrimination and Efficiency in Education. *International Economic Review* **46**: 1009-1031.

[19] De Fraja, G. and Iossa, E. 2002 Competition among Institutions and the Emergence of the Elite University. *Bulletin of Economic Research* **54**: 275-293.

[20] De Fraja, G. and Landeras, P. 2006 Could Do Better: The Effectiveness of Incentives and Competition in Schools. *Journal of Public Economics* **90**: 189-213.

[21] De Fraja, G., Oliveira, T. and Zanchi, L. 2010. Must Try Harder. Evaluating the Role of Effort on Examination Results. *Review of Economics and Statistics* **92**: 577-597.

[22] De Groot, M.H. 1970. *Optimal Statistical Decisions*. McGraw-Hill. New York.

[23] Dewatripont M., Jewitt I. and Tirole, J. 1999a. The Economics of Career Concerns, Part I: Comparing Information Structures. *Review of Economic Studies* **66**: 183-198.

[24] Dewatripont M., Jewitt I. and Tirole, J. 1999b. The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies. *Review of Economic Studies* **66**: 199-217.

[25] DfES. 2003. The Future of Higher Education. Department for Education and Skills, HMSO 5735.

[26] Downey, D.B. and von Hippel, P.T. and Beckett A. Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year. *American Sociological Review* **69**: 613-635.

[27] Edin, P.A. and Lagerström, J. 2006. Blind Dates: Quasi Experimental Evidence in Discrimination. Institute for Evaluation of Labour Market and Education Policy, Uppsala 2006:4.

[28] Eells, K., Davis, A., Havighurst, R.J., Herrick, V.E. and Tyler, R.1951. *Intelligence and Cultural Differences*. Chicago University Press.

[29] Epple, D., Romano, R., Sarpca, S. and Sieg, H. 2006. Profiling in Bargaining over College Tuition. *Economic Journal* **11**: 6459-479.

[30] Epple, D., Romano, R., and Sieg, H. 2003. Peer Effects Financial Aid and Selection of Students to Colleges and Universities: An Empirical Analysis. *Journal of Applied Econometrics* **18**: 501-525.

[31] Epple, D., Romano, R., and Sieg, H. 2008. Diversity and Affirmative Action in Higher Education. *Journal of Public Economic Theory* **10**: 475-501.

[32] Farber, H.S. and Gibbons, R. 1996. Learning and Wage Dynamics. *Quarterly Journal of Economics* **111**: 1007-1047.

[33] Galindo-Rueda, F. and Vignoles, A. 2003 Class Ridden or Meritocratic? An Economic Analysis of Recent Changes in Britain. Centre for the Economics of Education, London School of Economics 32.

[34] Gary-Bobo, R. and Trannoy, A. 2008. Efficient Tuition Fees and Examinations. *Journal of the European Economic Association* **6**: 1211–1243.

[35] Goldin, C. and Rouse, C. 2000 Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *American Economic Review* **90**: 715-741.

[36] Hauser, R.M. 2002. Meritocracy, Cognitive Ability, and the Sources of Occupational Success. Center for Demography and Ecology, University of Wisconsin-Madison 32.

[37] Lee, V.E. and Burkam, D.T. 2002. Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School. Economic Policy Institute, Washington.

[38] Lundberg, S.J. 1991 The Enforcement of Equal Opportunity Laws under Imperfect Information: Affirmative Action and Alternatives. *Quarterly Journal of Economics* **106**: 309-326.

[39] Lundberg, S.J. and Startz, R. 1983. Private Discrimination and Social Intervention in Competitive Labor Market American Economic Review73: 340-347.

[40] MacLeod, W. Bentley, and Miguel Urquiola. 2009. Anti-Lemons: School Reputation and Educational Quality. National Bureau of Economic Research Working Paper 15112.

[41] MacLeod, W.B. and Urquiola, M. 2015. Reputation and School Competition. American Economic Review 105: 3471-3488.

[42] Marton, F., Hounsell, D. and Entwistle, N. 1997. The Experience of Learning. Scottish Academic Press. Edinburgh.

[43] Mayer, S.E. 1997. What Money Can't Buy: Family Income and Children's Life Chances. Harvard University Press. Cambridge.

[44] Meghir, C., Palme, M. and Simeonova, E. 2012. How Does Education Affect Labor Market Outcomes: The Effect of Schooling on Cognitive and Non-Cognitive Skills. mimeo

[45] Morgan, P. 2009. Risk Factors for Learning-Related Behavior Problems at 24 Months of Age: Population-Based Estimates. Journal of Abnormal Child Psychology 37: 401-413.

[46] Moro, A. and Norman, P. 2003. Affirmative Action in a Competitive Economy. Journal of Public Economics 87: 567-594.

[47] Moro, A. and Norman, P. 2004. A General Equilibrium Model of Statistical Discrimination. Journal of Economic Theory 114: 1-30.

[48] New York Times, 2012a. Rethinking Affirmative Action. http://www.nytimes.com/2012/10/14/sunday-review/rethinking-affirmative-action.html?

[49] New York Times, 2012b. Affirmative Action - A Complicated Issue for Asian-Americans. http://www.nytimes.com/2012/11/04/education/edlife/affirmative-action-a-complicated-issue-for-asian-americans.html?

[50] New York Times, 2016. Missing

[51] Norman, P. 2003. Statistical Discrimination and Efficiency. Review of Economic Studies 70: 615-627.

[52] Pacelli, K. A. 2011. Fisher v. University of Texas at Austin: Navigating the Narrows between Grutter and Parents Involved. Maine Law Review. 63 (2): 569-592.

[53] Phelps, E. 1972. The Statistical Theory of Racism and Sexism. American Economic Review 62: 659-661.

[54] Phillips, M., Brooks-Gunn, J., Duncan, G.J. and Klebanov, P.K. 1998. Family Background, Parenting Practices, and the Black-White Test Score Gap. in Charles Jencks and Meredith Phillips (eds) The Black-White Test Score Gap. The Brookings Institute. Washington.

[55] Reardon, S. 2003. The Growth of Racial - Ethnic and Socioeconomic Test Score Gaps in Kindergarten and First Grade. Paper presented at the Annual Conference of the American Sociological Association.

[56] Rose, H. 2006. Do gains in test scores explain labor market outcomes? Economics of Education Review 25: 430-446.

[57] Spence, M.A. 1973. Job Market Signalling. Quarterly Journal of Economics 873: 55-74.

[58] Warren, J.R., Grodsky, E. and Lee, J.C. 2008. State High School Exit Examinations and Postsecondary Labor Market Outcomes Sociology of Education 81: 77–107.

[59] Sowell, T. 2004. Affirmative Action Around the World: An Empirical Study. Yale University Press.

Appendix

Proof of Proposition 2

The school problem is

$$\max_{b_1,\ldots,b_S} \sum_s g_s \left(\frac{\rho_t}{\rho_s+\rho_t} \left(\bar{a}_s + e_s^* + \lambda\left(b_s\right)\right) + \frac{\rho_s}{\rho_s+\rho_t} \left(\bar{a}_s + \eta e_s^* + \beta\lambda\left(b_s\right)\right) \right) +$$
$$\mu \sum_s g_s m_s \left(\bar{a}_s + e_s^* + \lambda\left(b_s\right)\right)$$
$$\text{s.t. } B + \sum_s g_s m_s \left(\bar{a}_s + e_s^* + \lambda\left(b_s\right)\right) \geqslant c\left(\sum_s g_s b_s\right),\ b_s \geqslant \bar{b}.$$

The first order conditions in problem (3) are:

$$\left(\frac{\rho_t+\beta\rho_s}{\rho_s+\rho_t} + \left(\mu+\nu\right) m_s\right) \lambda'\left(b_s\right) = \nu c'\left(\sum_s g_s b_s\right), \qquad s = 1, ..., S, \tag{27}$$

where ν is the Lagrange multiplier of the budget constraint. Since the RHS is constant in s, it is clear that the LHS in the above must be constant across social groups. Total differentiation of the LHS of the above gives:

$$\begin{aligned}&\left(\frac{\rho_t+\beta\rho_s}{\rho_s+\rho_t} + \left(\mu+\nu\right) m_s\right) \lambda''\left(b_s\right) db_s + \\ &\lambda'\left(b_s\right)\left(\frac{\left(\beta-1\right)\rho_t}{\left(\rho_s+\rho_t\right)^2}\rho'\left(s\right) + \left(\mu+\nu\right) m'\left(s\right)\right) ds \\ &= 0\end{aligned} \tag{28}$$

Derive ν from (27) to obtain that:

$$\mu + \nu = \frac{\frac{\rho_t+\beta\rho_s}{\rho_s+\rho_t} + \mu\frac{c'(\cdot)}{\lambda'(b_s)}}{\frac{c'(\cdot)}{\lambda'(b_s)} - m_s},$$

and put it into (28):

$$\frac{db_s}{ds} = -\frac{\lambda'\left(b_s\right)}{\lambda''\left(b_s\right)} \frac{\left(1 - \frac{\lambda'(b_s)}{c'(\cdot)} m_s\right)\frac{(\beta-1)\rho_t}{(\rho_s+\rho_t)^2}\rho'\left(s\right) + \left(\frac{\lambda'(b_s)}{c'(\cdot)}\frac{\rho_t+\beta\rho_s}{\rho_s+\rho_t} + \mu\right) m'\left(s\right)}{\frac{\rho_t+\beta\rho_s}{\rho_s+\rho_t} + \mu m_s}$$

then rearrange to obtain (13).

Proof of Corollary 2

Differentiation of (20) with respect to s yields

$$\frac{d\left(\bar{a}_s+\frac{\eta\rho_s+\rho_t}{\rho_s+\rho_t}e_s+\frac{\beta\rho_s+\rho_t}{\rho_s+\rho_t}\lambda\left(b_s\right)\right)}{ds}=\bar{a}_s'+\frac{d}{ds}\left(\frac{\eta\rho_s+\rho_t}{\rho_s+\rho_t}e_s\right)+\frac{d}{ds}\left(\frac{\beta\rho_s+\rho_t}{\rho_s+\rho_t}\lambda\left(b_s\right)\right),$$

$$\frac{d\left(\bar{a}_s+\frac{\eta\rho_s+\rho_t}{\rho_s+\rho_t}e_s+\frac{\beta\rho_s+\rho_t}{\rho_s+\rho_t}\lambda\left(b_s\right)\right)}{ds}=\bar{a}_s'+\frac{e_s'(\rho_t+\rho_s)(\rho_t+\eta\rho_s)-(1-\eta)\rho_t e_s\rho_s'}{(\rho_t+\rho_s)^2}+\frac{\frac{db_s}{ds}(\rho_t+\rho_s)\lambda'(b_s)(\rho_t+\beta\rho_s)-(1-\beta)\rho_t\lambda(b_s)\rho_s'}{(\rho_t+\rho_s)^2}.$$

We rearrange it to collect $\frac{db_s}{ds}$,

$$\begin{aligned}\frac{d\left(\bar{a}_s+\frac{\eta\rho_s+\rho_t}{\rho_s+\rho_t}e_s+\frac{\beta\rho_s+\rho_t}{\rho_s+\rho_t}\lambda\left(b_s\right)\right)}{ds}=&\\ &\bar{a}_s'-\left(\rho_s+\rho_t\right)^{-2}\left[\rho_s'\rho_t\left(e_s\left(1-\eta\right)+\lambda\left(b_s\left(s\right)\right)\left(1-\beta\right)\right)\right]+\\ &\frac{\eta\rho_s+\rho_t}{\rho_s+\rho_t}\frac{\rho_t\rho'\left(s\right)}{\psi''\left(e_s^*\right)\left(\rho\left(s\right)+\rho_t\right)^2}\\ &-\frac{\beta\rho_s+\rho_t}{\rho_s+\rho_t}\lambda'\left(b_s\right)\frac{db_s}{ds}.\end{aligned}$$

Substituting (13), we get the result in the corollary.

Proof of Proposition 4

The first order conditions in problem (25) are:

$$\left(1+\left(\mu+\nu\right)m_s\right)\lambda'\left(b_s\right)=\nu c'\left(\sum_s g_s b_s\right),\qquad s=1,...,S,$$

Total differentiation of the LHS of the above gives:

$$\left(1+\left(\mu+\nu\right)m_s\right)\lambda''\left(b_s\right)db_s+\lambda'\left(b_s\right)\left(\mu+\nu\right)m'\left(s\right)ds=0 \tag{29}$$

Derive ν from (27) to obtain that:

$$\mu+\nu=\frac{1+\mu\frac{c'(\cdot)}{\lambda'(b_s)}}{\frac{c'(\cdot)}{\lambda'(b_s)}-m_s}$$

and put it into (29):

$$\left(1+\frac{1+\mu\frac{c'(\cdot)}{\lambda'(b_s)}}{\frac{c'(\cdot)}{\lambda'(b_s)}-m_s}m_s\right)\lambda''\left(b_s\right)db_s+\lambda'\left(b_s\right)\frac{1+\mu\frac{c'(\cdot)}{\lambda'(b_s)}}{\frac{c'(\cdot)}{\lambda'(b_s)}-m_s}m'\left(s\right)ds=0,$$

$$\left(\left(\frac{c'\left(\cdot\right)}{\lambda'\left(b_s\right)}-m_s\right)+\left(1+\mu\frac{c'\left(\cdot\right)}{\lambda'\left(b_s\right)}\right)m_s\right)\lambda''\left(b_s\right)db_s+\lambda'\left(b_s\right)\left(1+\mu\frac{c'\left(\cdot\right)}{\lambda'\left(b_s\right)}\right)m'\left(s\right)ds=0,$$

$$\left(\frac{c'\left(\cdot\right)}{\lambda'\left(b_s\right)}\left(1+\mu m_s\right)\right)\lambda''\left(b_s\right)db_s+\lambda'\left(b_s\right)\left(1+\mu\frac{c'\left(\cdot\right)}{\lambda'\left(b_s\right)}\right)m'\left(s\right)ds=0,$$

$$\frac{db_s}{ds}=-\frac{\lambda'\left(b_s\right)}{\lambda''\left(b_s\right)}\frac{\left(1+\mu\frac{c'(\cdot)}{\lambda'(b_s)}\right)m'\left(s\right)}{\frac{c'(\cdot)}{\lambda'(b_s)}\left(1+\mu m_s\right)}.$$

Rearrange to obtain (26).