Social Background in School Attainment and Job Market

Alessandro Tempieri, University of Leicester, UK

Working Paper No. 09/26
December 2009
Social Background in School Attainment and Job Market*

Alessandro Tampieri†
University of Leicester

December 7, 2009

Abstract

This paper studies how social background affects schooling attainment and job opportunities from a theoretical perspective. We analyse the interaction between a school and an employer when students attend school and then go to the job market. Students differ in ability and belong to different social groups. Our results suggest that the employer makes use of social background in the recruitment decisions: this favours advantaged students, as they are more likely to have high ability. In turn, the school optimally provides disadvantaged students with less teaching than advantaged students, given the same level of ability.

JEL Numbers: C73, I21, J24

Keywords: Social background, Ability.

*I would like to thank Gianni De Fraja for his continuous support and encouragement. I am also grateful to Javier Rivas, Gaia Garino, Piercarlo Zanchettin, Matteo Lippi Bruni, Piero Pasotti for helpful comments. A previous version of this paper has been presented to the Scottish Economic Society Annual Conference, 2009. All errors are mine.

†University of Leicester, Department of Economics, Leicester, LE1 7RH, UK, University of Rome “La Sapienza”, Department of Public Economics, Via del Castro Laurenziano 6, 00161 Rome, Italy; email: at198@le.ac.uk.
1 Introduction

There is substantial evidence that social background influences educational results and job opportunities\(^1\). Although these empirical findings, few contributions have tried to understand the mechanism behind it.

This paper examines the impact of social background in educational attainment and job opportunities by highlighting the interaction between schools and employers. We assume that job opportunities are influenced by social background directly, as employers use this information as a signal of ability, and indirectly, as the schools’ grading strategy takes into account the employers’ hiring decisions.

We study the equilibria of a game between a school and an employer where the school decides the effort in teaching students and the employer evaluates whether to hire them or not. Students can have high or low ability and advantaged or disadvantaged background.

There are two major assumptions in the paper. First, while both agents know the students’ social background, the school knows students’ ability but the employer uses grades as signals of ability. Second, the distribution of ability among individuals is different between communities. Disadvantaged individuals are less likely to have high ability. Research has documented that family and environmental factors are major predictors of cognitive abilities (Cunha \textit{et al.}, 2006, Carneiro and Heckman, 2003, Joshi and McCulloc, 2000). These differences in the distribution of abilities may be due to a variety of causes: differences in social skills, in the distribution of cognitive skills in the different groups, and so on. Environments that do not stimulate the young and fail to cultivate cognitive skills place children at disadvantage in terms of ability\(^2\). Advantaged students grow in a more stimulating environment, which enhances the students’ ambitions, the language talked at home is finer, the topics of conversation are closer to the schools’ disciplines, and so on.

\(^1\)For educational attainment, see Haveman and Wolfe (1995) for a discussion, while Galindo-Rueda and Vignoles (2005) and Marcenaro-Gutierrez \textit{et al.} (2007) give some recent contributions. For job opportunities, Glyn and Salverda (2000) and Berthoud and Blekesaune (2006) show how a disadvantaged background affects the chances of finding a job in OECD countries and UK, respectively.

\(^2\)An alternative explanation is that disadvantage arises from lack of financial resources (Carneiro and Heckman, 2003).
Our results suggest that, education being an imprecise signal of ability, the employer makes use of social background in their hiring decisions, by favouring, *ceteris paribus*, advantaged students, as they are more likely to have high ability. In particular, the employer evaluates firstly advantaged students and, only once finished their labour supply, disadvantaged students. Given the employer’s strategy, the school’s optimal response is to provide disadvantaged students with less teaching, given the same level of ability. This is in contrast to De Fraja (2005), by which, for efficiency purposes, disadvantaged and high-ability students should receive more education than advantaged and high-ability students.

In economics of education, differences in social background have been introduced in the analysis of school policy reforms. Epple and Romano (1998) study the effects of tuition vouchers and peer-groups effects in a competition between private and public schools. In their model, students differ in ability and income. In equilibrium, advantaged and low-ability students pay a tuition premium while disadvantaged and high-ability students receive a tuition discount. Nechyba (1999) analyses model with different public school communities in presence of private schools. Peer effects, which are correlated with the socioeconomic status of households, influence the students’ attainment. The introduction of vouchers in presence of migration patterns would benefit public schools in poor communities while hurting public schools in wealthy communities. Unlike these contributions, we choose not to consider the competition between public and private schools and focus on the impact of social background.

Also, the schools’ evaluation in our model may be interpreted as the university’s admission policy: in this context, our analysis may be linked to the theoretical literature on universities’ admission strategies. This literature follows:

---

3Although these related results, De Fraja (2005) investigates the impact of antidiscrimination policies.

4Epple et al. (2006) model the college admission as a bargaining process between students and universities under asymmetric information. Gary-Bobo and Trannoy (2007) study the optimal policy admission as a mix of tuition fees and admission tests in order to solve the two-sided information asymmetry about students ability (both universities and students do not know it exactly). Fernandez and Gali' (1999) compare the performance of markets and tournaments as allocative mechanisms. They consider a matching between students and universities when borrowing constraints are present. The presence of the latter makes the tournament the most efficient mechanism.
cuses on asymmetric information about the student ability, and disregards the students’ social background. Our contribution to this strand can be the introduction of social background when schools evaluate students.

The remainder of the paper is organised as follows. The model is presented in Section 2. Section 3 examines the equilibria in the case with no differences in social background. This is introduced in Section 4 and a possible government intervention is discussed in Section 5. Section 6 concludes.

2 The model

We study the interaction between a single school and a single employer. This occurs as both serve a number of students, with measure normalised to one. In order to have as simple a framework as possible, we study abstract from factors such as competition between schools and between employers.

2.1 School

The school prepares its students for the final exam. The possible exam’s outcomes are only two, a low or a high grade.

The school knows the students’ ability. This informational power can be motivated by understanding schooling signals, such as tests along with school attendance. Students can have high ability ($\theta_H$) or low ability ($\theta_L$). In our model, a high-ability student has more chance to perform well at school and at work. Let $p \in [0, 1]$ denote the probability that a student has high ability.

The school has resources in order to prepare students. As resources, we consider a variety of aspects: the quantity of teaching, the quality of buildings and classroom equipment, the teachers’ effort in teaching, and so on. From now on, we refer to all these aspects as “teaching”. In addition to “basic” teaching, the school can provide some students with additional resources, extra tuition, trips and so on. If the school does so, the student is more likely to obtain a...
high grade in the final exam. The school pays a fixed amount $c > 0$ for each student who receives extra teaching.

Here we simplify by assuming that, if the school provides extra teaching to a high-ability student, the student’s probability of obtaining a high grade in the final exam is equal to one, otherwise it is equal to an exogenous probability $\eta \in (0, 1)$. On the other hand, in the case that the school provides extra teaching to a low-ability student, her probability to obtain a high grade is $\eta \in (0, 1)$ and zero otherwise\(^7\).

We assume that, in the case where a student is hired by the employer, the school obtains a benefit $\mu > 0$. In fact the school increases its reputation and credibility if its students find a job after the school time. Finally, the school’s payoff is given by the difference between benefit and cost.

### 2.2 Employer

The employer decides whether or not to hire a student, and offers a single type of job. The employer can only observe the student’s grade, and use it as a signal of ability.

For the sake of simplicity, we rule out uncertainty in the market where the employer operates and we assume that the students’ ability determines the employer’s profit entirely. The employer’s profit is normalised to $\nu > 0$ if the student has high ability and to $-1$ if the student has low ability. The assumption captures the fact that low-ability employees generally have a marginal productivity which is lower than salary cost.

We denote labour demand as $\Phi \in (0, 1)$. Note that labour demand cannot satisfy the entire labour supply, as the total mass of students is one.

### 2.3 The Game between school and employer

The interaction can be described as a three-stage game of the following form (Figure 1). At the first stage, nature draws a student type.

---

\(^7\)The two events “high-ability with no extra teaching obtains a low grade” and “low-ability with extra teaching obtains a high grade” have the same probability to occur in order to simplify the algebra. To give to these two events a different probability would just complicate the analysis without adding any insight.
At the second stage, the school chooses whether to provide the student with extra teaching. At the end of the second stage, the student attends school and then takes the final exam\(^8\). At the third stage, the student applies for a job. The employer, who does not know the student’s ability, decides whether to hire her. In order to come to a decision, the employer observes the grade obtained by the student in the final exam. Moreover, the employer knows the probability that a student has high ability and also of the structure of the school’s cost-benefit.

### 3 Equilibria

We study the perfect Bayesian equilibria of this game. In this economy, a perfect Bayesian equilibrium is a combination of school’s and employer’s strategies and beliefs given that the employer does not know the student’s type.

\(^8\)For simplicity, we assume away the influence of student’s effort.
We start by making the following assumptions.

**Assumption 1**

(i) \( p + (1 - p) \eta \leq \Phi \).

(ii) \( \mu > \max\left\{ \frac{c}{\eta}, \frac{c}{1 - \eta} \right\} \).

These assumptions rule out unrealistic equilibria and give to the model consistency with the empirical evidence. Without these assumptions, the school’s and employer’s interest coincide: equilibria occur where education signals perfectly ability to the employer, as the school would never find convenient to provide extra teaching to low-ability students. Assumption 1.(i) says that the amount of high-degree students is lower than the labour demand. This assumption has two purposes. First, it implies that the school cannot maximise its payoff by providing extra teaching only to high ability students. Second, it rules out equilibria which are consistent with our results but do not add any insight. Assumption 1.(ii) rules out that the school’s benefit obtained by low-ability students is not high enough compared to the cost of extra teaching, \( c \). Otherwise every high-grade student would have high ability and the employer would have no problems for recognising them.

After introducing these assumptions, there are two possible equilibria, defined as follows.

**Definition 1.** *High-employment equilibrium.*

(i) The school provides extra teaching to a high-ability student with probability 1 and to low-ability student with probability \( \frac{\Phi - p}{1 - p} \).

(ii) The employer hires a high-grade student with probability 1 and a low-grade student with probability 0.

**Definition 2.** *Low-employment equilibrium.*

(i) The school provides extra teaching to each high-ability student and to a low-ability student with probability \( \frac{\Phi - p}{1 - p} \frac{p}{1 - p} \eta \) \( \in (0, 1) \).

(ii) The employer hires a high-grade student with probability \( \frac{\Phi}{\mu \eta} \in (0, \Phi) \) and none of low-grade students.

The following proposition shows which equilibrium occurs according to the value of \( p \).
**Proposition 1** Let Assumption 1 hold. The high-employment equilibrium occurs if \( p > \frac{\eta}{\nu + \eta} \); the low-employment equilibrium occurs if \( p \leq \frac{\eta}{\nu + \eta} \).

**Proof.** See Appendix. ■

Figure 2 illustrates Proposition 1. The horizontal axis is \( p \) while the vertical axis is \( \eta \). The diagonal line is where \( \Phi = p + (1 - p) \eta \). By Assumption 1.(i), we study the equilibria below this line. The upward-sloping line represents the critical point where \( \frac{\eta}{\nu + \eta} \).

The **high-employment equilibrium** occurs if the probability that a student has high ability is greater than \( \frac{\eta}{\nu + \eta} \). Since it is very likely that a student has high ability, the employer maximises his profit by hiring as many high-grade students as he can. Given this strategy, the school provides extra teaching to every high-ability student, and also to a number of low-ability students which fills the labour demand, that is \( \frac{\eta}{\nu + \eta} \).

As \( p \) fall to \( \frac{\eta}{\nu + \eta} \), the **low-employment equilibrium** occurs. Since \( p \) is not high enough, to hire \( \Phi \) high-grade student would worse off the employer’s profit: the best strategy now is to hire high-grade students only with a certain probability lower than \( \Phi \) which maximises the expected profit. As the number of hired high-grade student diminishes, the school provides extra teaching to a low-
ability student only with a certain probability lower than \( \frac{p - p_1}{1 - p} \). In this case, education works better as a signalling device of ability compared to the previous equilibrium, as the probability that a high-grade student has also high ability is larger.

4 Social background

In this section we introduce social background. Suppose that a student can have either advantaged (a) or disadvantaged (d) social background\(^9\), which is public information. We denote as \( \lambda \in [0, 1] \) the proportion of advantaged students. Social background affects the chance of having high ability. Let \( p_d, p_a \in [0, 1] \) be the probability that a disadvantaged or advantaged student has high-ability, respectively. Also, we assume\(^{10} \) that \( p_a > p_d \).

We assume that advantaged and disadvantaged students live in separated communities. In each community there is one school. We assume that there is no competition between the two schools: any group of student who lives in a certain community must attend the local school. Schools are completely identical in terms of formative capacity and technology.

The timing of the game is modified as follows. In the first stage, nature allocates a student either in the advantaged or disadvantaged community. Also, nature draws the ability type of the student. In second stage, the advantaged and the disadvantaged schools have the same actions as the unique school in the previous case. In the third stage, the employer decides whether to hire the student like in the previous case, but now he has informations about the student’s social background. Hence his actions can be “hire” or “not hire” an advantaged and-high-grade student, an advantaged and low-grade student, a disadvantaged and high-grade student, or a disadvantaged and low-grade student.

In order to show the equilibria with social background, we need to make different technical assumptions.

\(^9\)This can be interpreted as a unique measure of family environment, peer groups (students learn better if they are in a group of abler students), neighbourhood and so on.

\(^{10}\)Note that, as both agents are aware of the students’ social background and probability of being able, it follows that they know the amount of \( p_a \) and \( p_d \).
Assumption 2 (i) $\Phi \in (\lambda (p_a + (1 - p_a) \eta) + (1 - \lambda) p_d, \\
\lambda (p_a + (1 - p_a) \eta) + (1 - \lambda) (p_d + (1 - p_d) \eta)$.
(ii) $\mu > \max \left\{ \frac{c}{\eta}, \frac{c}{1-\eta} \right\}$.

Assumption 2.(i) links the total labour demand between the proportion of advantaged and high-grade students and the proportion of all high-grade students. This assumption allows us to highlight those equilibria where the presence of social background is interesting\textsuperscript{11}. If the upper bound would not hold, there is job for every graduate student, the employer would not be forced to choose between advantaged and disadvantaged students and hence the influence of social background would not be striking. If the lower bound does not hold, equilibria would exist where disadvantaged students are simply ignored by the employer as he will consider only advantaged students to hire. This occurs as the expected payoff given by advantaged students is, ceteris paribus, higher than the expected payoff given by disadvantaged students. Assumption 2.(ii) is equivalent to Assumption 1.(ii) with social background.

After introducing these assumptions, the possible equilibria defined in this setting are the following:

Definition 3. **High-employment equilibrium.** (i) The advantaged school gives extra teaching to each student with probability 1; the disadvantaged school gives extra teaching to a high-ability student with probability 1 and to a low-ability student with probability $\frac{\Phi - \lambda (p_a + \eta (1-p_a)) - (1-\lambda) p_d}{p_d + \eta (1-p_d)}$.

(ii) The employer hires an advantaged and high-grade student with probability 1, an advantaged and low-grade with probability 0, a disadvantaged and high-grade with probability $\frac{\Phi - \lambda (p_a + \eta (1-p_a))}{p_d + \eta (1-p_d)}$ and a disadvantaged and low-grade with probability 0.

Definition 4. **Middle-employment equilibrium.** (i) The advantaged school gives extra teaching to each student with probability 1; the disadvantaged school gives extra teaching to a high-ability student with probability 1 and to a low-ability with probability $\frac{\Phi - \lambda (p_a + \eta (1-p_a)) - (1-\lambda) p_d}{p_d + \eta (1-p_d)}$.

(ii) The employer hires an advantaged and high grade student with probability 1, an advantaged and low-grade with probability 0, a disadvantaged and high-grade with probability 1, an advantaged and low-grade with probability 0, a disadvantaged and

\textsuperscript{11}The equilibria not presented are, nonetheless, consistent with our results. Under request, we can provide the equilibria on the entire parameter space.
high-grade with probability $\frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \frac{c}{\mu \eta}$ and a disadvantaged and low grade with probability 0.

**Definition 5. Low-employment equilibrium.** (i) The advantaged school gives extra teaching to a high-ability student with probability $\frac{p_a}{(1-p_a) \eta}$; the disadvantaged school gives extra teaching to a high-ability student with probability 1 and to a low-ability with probability

$$\frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \frac{p_d}{(1-p_a) \eta}.$$ 

(ii) The employer hires an advantaged and high-grade student with probability $\frac{c}{\mu \eta}$, an advantaged and low-grade with probability 0, a disadvantaged and high-grade with probability $\frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \frac{c}{\mu \eta}$ and a disadvantaged and low-grade with probability 0.

The following proposition shows when equilibria occur according to the values of $p_a$ and $p_d$.

**Proposition 2** Let Assumption 2 hold. The high-employment equilibrium occurs if $p_a > \frac{\eta}{\nu + \eta}$ and $p_d > \frac{\eta}{\nu + \eta}$; the middle-employment equilibrium occurs if $p_a > \frac{\eta}{\nu + \eta}$ and $p_d \leq \frac{\eta}{\nu + \eta}$; the low-employment equilibrium occurs if $p_a \leq \frac{\eta}{\nu + \eta}$ and $p_d \leq \frac{\eta}{\nu + \eta}$.

**Proof.** See Appendix. ■

Figure 3 illustrates Propositon 2. The higher diagonal line is where $\Phi = \lambda (p_a + (1-p_a) \eta) + (1-\lambda) p_d$. The lower diagonal line is where:

$$\Phi = \lambda (p_a + (1-p_a) \eta) + (1-\lambda) (p_d + (1-p_d) \eta).$$

By Assumption 2.(i), we study the equilibria between these lines. The continuous and the dashed upward-sloping lines represent the critical point of $p_a$ and $p_d$, respectively. In the area on the left of the dashed line, the low-employment equilibrium occurs, between the lines the middle-employment equilibrium occurs and on the right of the continuous line the high-employment equilibrium occurs.

To interpret Proposition 2, begin by looking at the employer' choice. The key assumption is $p_a > p_d$, which makes the employer to obtain, *ceteris paribus,*
a higher expected payoff by hiring advantaged students. Therefore the disadvantaged students opportunity of being hired is only subsequent to the recruitment within the advantaged community.

Second, in every scenario, the presence of advantaged students affects the job opportunities of disadvantaged students. This can be easily seen by increasing $\lambda$: regardless of the equilibrium, the probability of being hired for a high-grade and disadvantaged student diminishes. Also, if we allow the labour demand to be lower than the amount of advantaged and high-degree students (by relaxing Assumption 2.(i)), then none of the disadvantaged students will be hired.

Third, the value of both $p_a$ and $p_d$ matters within each community. If $p_a$ is low, the employer would hire an amount of high-grade and advantaged student lower than the number available. If $p_d$ is low, the employer would hire an amount of high-grade and disadvantaged student lower than the residual labour demand.

The employer’s strategy in turn affects the schools policy. The advantaged school does not have to take into account the labour demand in order to plan its strategy, since the assumption $p_a > p_d$ makes its students to be “privileged” by
the employer. This is also ensured by Assumption 2.(i), by which the amount of high-grade and advantaged student is lower than the labour demand. Instead the disadvantaged school, in order to maximise its payoff, should not consider providing extra teaching to the students of an amount greater than the difference between the labour demand and the amount of high-degree and advantaged students. Hence the optimal response of the disadvantaged school to the employer’s strategy is to provide less education than the one provided by the advantaged school, given the same level of students’ ability. This result is sharply in contrast with De Fraja (2005): here, efficiency requires that disadvantaged and high-ability students should receive more education compared to high-ability and advantaged students.

5 Concluding remarks

This paper studies how social background may affect schooling attainment and employer’s hiring behaviour in the job market. We study the interaction between a school and an employer when students differ in ability and belong to different social groups. According to our results, employers make also use of candidates’ social background as a signal of ability, by favouring advantaged students as they are more likely to have high ability. In turn, the school optimally provides less teaching to disadvantaged students, given the same level of ability.

This approach can be extended in several ways. First, the problem can be seen from a normative perspective. A government intervention could mitigate the issue, by assigning education standards that maximise the disadvantaged welfare function. Second it seems natural to consider different schools for each social group and see how school competitions modifies the schools’ teaching strategy. Finally, it would be interesting to examine this framework along different generations in order to explain segregation mechanisms. The analysis of an extended model in these directions is left for future work.
References


Appendix

Proof of Proposition 1

Let us assume that \( x_L, x_H \in [0, 1] \) are the probabilities that the school provides with extra teaching a low-ability or high-ability student, respectively, while \( z_U, z_D \in [0, 1] \) are the probabilities that the employer hires a student with a high (\( U \)) or low (\( D \)) grade. Note that \( -1 \leq \nu > 0 \) are the employer’s profit if the hired student has low or high ability, respectively, while \( \mu \) is the school’s benefit if a student is hired. Finally, \( c \) is the school’s cost of providing extra teaching.

The equilibrium concept necessary to solve this game is the perfect Bayesian equilibrium. A perfect Bayesian equilibrium necessitates the following requirements (Gibbons, 1992):

**Requirement 1.** After observing a grade \( g_j \) where \( j \in \{U, D\} \), the employer must have a belief about which type of student could have obtained it. We denote this belief by the probability distribution \( \pi (\theta_i \mid g_j) \), where \( i \in \{H, L\} \).

**Requirement 2.** For each grade, the employer’s action \{hire, not to hire\} must maximise his expected profit, given the belief \( \pi (\theta_i \mid g_j) \).

**Requirement 3.** Since the school has complete information and moves before the employer, its strategy must maximise its payoff, given the employer’s strategy.
Requirement 4. For each $g_j$, if there exists $\theta_i$ such that $g_j^*(\theta_i) = g_j$, then the employer’s belief must follows from Bayes’ rule and the school’s strategy.

A perfect Bayesian equilibrium is a pair of school and employer’s strategies and a belief satisfying the Requirements 1,2,3 and 4. Along the proof, we consider separately each agent.

Case 1. $p > \frac{\eta}{\nu + \eta}$

Employer. With $p > \frac{\eta}{\nu + \eta}$, we assume that the employer’s strategies are $z_U = 1; z_D = 0$. According to the school’s strategy, in the case that the student has a high grade the employer’s beliefs are: $\pi(\theta_H | g_U) = \frac{p}{p + \eta(1-\nu)}$ and $\pi(\theta_L | g_U) = \frac{\eta(1-p)}{p + \eta(1-p)}$. Thus the expected profits\(^{12}\) are $\Pi^E_U = \frac{p}{p + \eta(1-\nu)} \nu - 1 \frac{\eta(1-p)}{p + \eta(1-p)}$ for hiring and $\Pi^N_U = 0$ for not hiring one high-grade student, respectively. We are assuming $z_U = 1$, so $\Pi^E_U > \Pi^N_U$ and therefore $\frac{p}{p + \eta(1-\nu)} \nu - 1 \frac{\eta(1-p)}{p + \eta(1-p)} > 0$. After few passages, we obtain $p > \frac{\eta}{\nu + \eta}$. By Assumption 1.(i), the amount of high-grade students is lower the labour demand, therefore this strategy occurs $p + \eta (1 - p)$ times, that is for each high-grade student who goes to the job market. Hence $z_U = 1$. According to the school’s strategy we are assuming, in the case that the student has a low grade the employer’s beliefs are: $\pi(\theta_H | g_D) = 0$ and $\pi(\theta_L | g_D) = 1$. Thus the expected profits\(^{13}\) are $\Pi^E_D = -1$ for hiring and $\Pi^N_D = 0$ for not hiring. We are assuming that $z_D = 0$, hence $\Pi^E_D < \Pi^N_D$, and it is always verified since $-1 < 0$.

School. With $p > \frac{\eta}{\nu + \eta}$, we assume that the school’s strategies are $x_L = \frac{\phi - p}{1 - p}; x_H = 1$. The school maximises its expected payoff given the employer’s strategies. In the case that the school teaches a high-ability student, the expected payoffs\(^{14}\) are $\Pi^T_H = \mu - c$ for giving extra teaching and $\Pi^{NT}_H = \mu \eta$ for not giving extra teaching. We are assuming that $\Pi^T_H > \Pi^{NT}_H$, that is $\mu - c > \mu \eta$, and therefore $\mu > \frac{c}{1 - \eta}$, which respects Assumption 1.(ii). By Assumption 1.(i),

\(^{12}\) The superscript of the employer’s expected profit indicates the strategy performed by the employer, where $E$ indicates “to hire” and $N$ not. The subscript specifies the student’s grade, where $U$ indicates a high grade and $D$ a low grade.

\(^{13}\) The superscript of the employer’s expected profit indicates the strategy performed by the employer, where $E$ indicates “to hire” and $N$ not. The subscript specifies the student’s grade, where $U$ indicates a high grade and $D$ a low grade.

\(^{14}\) The superscript of the school’s expected profit specifies the school’s strategy, where $T$ indicates “to provide extra teaching” and $NT$ not. The subscript indicates the student’s ability.
the amount of high-ability students is lower than the labour demand, therefore this strategy occurs \( p \) times, that is for each high-grade student who goes to the job market. Hence \( x_H = 1 \). For a number \( \frac{\Phi - p}{1 - p} \) of low-ability students\(^{15}\), the expected payoffs are \( \Pi^T_L = \mu \eta - c \) for giving extra teaching and \( \Pi^{NT}_L = 0 \) for not giving extra teaching. We are assuming that, for this amount of students, \( \Pi^T_H > \Pi^{NT}_H \), that is \( \mu \eta - c > 0 \) and \( \mu > \frac{\xi}{\eta} \), which respects Assumption 1.(ii).

For the remainder of low-ability students, the expected payoffs are \( \Pi^T_L = -c \) for giving extra teaching and \( \Pi^{NT}_L = 0 \) for not giving extra teaching. Hence the school never provides extra teaching to these students, as \( -c < 0 \).

**Case 2.** \( p \leq \frac{\eta}{\nu + \eta} \)

**Employer.** We assume that the employer’s strategies are \( z_U = \frac{c}{\nu} \); \( z_D = 0 \). According to the school’s strategy, in the case that the student has a high grade the employer’s beliefs are: \( \pi (\theta_H \mid g_U) = \frac{p}{p + \eta \xi L(1-p)} \) and \( \pi (\theta_L \mid g_U) = \frac{\eta \xi L(1-p)}{p + \eta \xi L(1-p)} \). Thus the expected profits are \( \Pi^E_U = \frac{p}{p + \eta} (1 - \nu) - 1 \frac{\eta (1-p)}{p + \eta (1-p)} \) for hiring and \( \Pi^N_U = 0 \) for not hiring one high-grade student. We are assuming that the employer randomises between hiring and not hiring, hence \( \Pi^E_U = \Pi^N_U \), therefore 

\[
\frac{p}{p + \eta \xi L(1-p)} (1 - \nu) - 1 \frac{\eta (1-p)}{p + \eta (1-p)} = 0.
\]

After few passages, we obtain \( x_L = \frac{p - \nu}{(1-p) \eta} \).

However, since the school never considers to teach a number of low-ability students greater than \( \Phi - p \) of the total amount of students (i.e. an amount \( \frac{\Phi - p}{1-p} \) of low-ability students), this probability has to be considered in the probability space \( \frac{\Phi - p}{1-p} \). Hence \( x_L = \frac{\Phi - p}{1-p} \frac{p - \nu}{(1-p) \eta} \). According to the school’s strategy we are assuming, in the case that the student has a low grade the employer’s beliefs are: \( \pi (\theta_H \mid g_D) = 0 \) and \( \pi (\theta_L \mid g_D) = 1 \). Thus the expected profits are \( \Pi^E_D = -1 \) for hiring and \( \Pi^N_D = 0 \) for not hiring. We are assuming that \( z_D = 0 \), hence \( \Pi^E_D < \Pi^N_D \), and it is always verified since \( -1 < 0 \).

**School.** We assume that the school’s strategies are \( x_L = \frac{\Phi - p}{1-p} \frac{p - \nu}{1-p \eta} \); \( x_H = 1 \). The school maximises its expected payoff given the employer’s strategies. In the case that the school teaches a high-ability student, the expected payoffs are \( \Pi^T_H = z_U \mu - c \) for giving extra teaching and \( \Pi^{NT}_H = z_U \mu \eta \) for not giving

\(^{15}\)The school teaches at most \( \Phi - p \) low-ability students. This in fact the maximum amount of low-ability students which give a positive payoff to the school. First, the school wants to teach high-ability students, since they give a higher payoff, and only residually low-ability students. If the school teaches a student more than \( \Phi - p \), this student cannot be hired for sure and hence the school is certain not to have a benefit from her.
extra teaching. We are assuming that $x_H = 1$, therefore $\Pi_H^T > \Pi_H^{NT}$, that is $z_U \mu - c > z_U \mu \eta$ and $\mu > \frac{c}{z_U (1-\eta)}$, which respects Assumption 1. $(ii)$. By Assumption 1. $(i)$, the amount of high-ability students is lower than the labour demand, therefore this strategy occurs $p$ times, that is for each high-grade student who goes to the job market. Hence $x_H = 1$. For a number $\frac{p - p}{1-p}$ of low-ability students, the expected payoffs are $\Pi_L^T = z_U \mu \eta - c$ for giving extra teaching and $\Pi_L^{NT} = 0$ for not giving extra teaching. The school randomises with the provisions of extra teaching, hence $\Pi_H^T = \Pi_H^{NT}$, that is $z_U \mu \eta - c = 0$ and $z_U = \frac{c}{\eta}$. By Assumption 1. $(i)$, the amount of high-grade students is lower the labour demand, therefore this strategy occurs $p + \eta (1 - p)$ times, that is for each high-grade student who goes to the job market. Hence $z_U = \frac{c}{\eta}$. For the remainder of low-ability students, the expected payoffs are $\Pi_L^T = c$ for giving extra teaching and $\Pi_L^{NT} = 0$ for not giving extra teaching. Hence the school never provides extra teaching to these students.

**Proof of Proposition 2**

As for Proposition 1 (see above) the equilibrium concept is the perfect Bayesian equilibrium. Let us assume that $x_{La}, x_{Ha} \in [0, 1]$ are the probabilities that the advantaged school provides extra teaching to a low-ability or high-ability student, respectively; $x_{Ld}, x_{Hd} \in [0, 1]$ are the probabilities that the disadvantaged school provides extra teaching to a low-ability or high-ability student, respectively; $z_{UA}, z_{DA} \in [0, 1]$ are the probabilities that the employer hires an advantaged student with a high or low grade; finally, $z_{Ud}, z_{Dd} \in [0, 1]$ are the probabilities that the employer hires a disadvantaged student with a high or low grade. Note, as a reminder, that $c$ and $c$ are the cost of providing extra teaching for the advantaged and disadvantaged school, respectively.

During the proof, we will proceed for showing the equilibria for the possible cases: $p_a > \frac{\eta}{\nu + \eta}$, $p_d > \frac{\eta}{\nu + \eta}$; $p_a > \frac{\eta}{\nu + \eta}$, $p_d \leq \frac{\eta}{\nu + \eta}$; $p_a \leq \frac{\eta}{\nu + \eta}$, $p_d \leq \frac{\eta}{\nu + \eta}$.

**Case 1.** $p_a > \frac{\eta}{\nu + \eta}$, $p_d > \frac{\eta}{\nu + \eta}$

**Employer.** We assume that the employer’s strategies are $z_{UA} = 1; z_{DA} = 0, z_{Ud} = \frac{\Phi - \lambda p_a + \eta (1-p_d)}{p_d + \eta (1-p_d)}; z_{Dd} = 0$. According to the schools’ strategy, in the case that an advantaged student has a high grade the employer’s beliefs are:
\[ \pi(\theta | g_U, a) = \frac{p_a}{p_a + \eta(1-p_a)} \] and \[ \pi(\theta | g_U, a) = \frac{\eta(1-p_a)}{p_a + \eta(1-p_a)}, \] while in the case that a disadvantaged student has a high grade, the employer’s beliefs are\[ \pi(\theta | g_U, d) = \frac{p_d}{p_d + \eta(1-p_d)} \] and \[ \pi(\theta | g_U, d) = \frac{\eta(1-p_d)}{p_d + \eta(1-p_d)}. \] Thus the expected profits\(^{16}\) are \( \Pi_U^E_a = \frac{p_a}{p_a + \eta(1-p_a)} \nu - 1 \frac{\eta(1-p_a)}{p_a + \eta(1-p_a)} \) for hiring and \( \Pi_U^N_a = 0 \) for not hiring one advantaged and high-grade student. On the other hand, the expected profits are \( \Pi_U^E_d = \frac{p_a}{p_a + \eta(1-p_a)} \nu - 1 \frac{\eta(1-p_a)}{p_a + \eta(1-p_a)} \) for hiring and \( \Pi_U^N_d = 0 \) for not hiring one disadvantaged and high-grade student. Since \( p_a > p_d \), the expected payoff obtained by hiring a high-grade and advantaged student is always higher than the expected payoff obtained by hiring a high-grade and disadvantaged student. Hence \( \Pi_U^E_a > \Pi_U^E_d \). Also, since we are assuming that \( \Pi_U^E_a > \Pi_U^N_a \), therefore \( \frac{p_a}{p_a + \eta(1-p_a)} \nu - 1 \frac{\eta(1-p_a)}{p_a + \eta(1-p_a)} > 0 \) and, after few passages, \( p_a > \frac{\eta}{\nu+\eta} \). This strategy occurs for \( \lambda(p_a + \eta(1-p_a)) \) times, that is the number of advantaged and high-grade students. Since, by Assumption 2.1(i), this is lower than the labour demand, we have that \( z_{ua} = 1 \). The remaining part of the labour demand comprehends \( \Phi - \lambda(p_a + \eta(1-p_a)) \) placements, that is there are placements for \( \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_a + \eta(1-p_a)} \) disadvantaged and high-grade students.

For this part of the labour demand, we are assuming that \( \Pi_U^E_d > \Pi_U^N_d \), therefore \( \frac{p_a}{p_a + \eta(1-p_a)} \nu - 1 \frac{\eta(1-p_a)}{p_a + \eta(1-p_a)} > 0 \) and, after few passages, \( p_d > \frac{\eta}{\nu+\eta} \). Since this strategy occurs for \( \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_a + \eta(1-p_a)} \) disadvantaged and high-grade students (i.e., the employer hires with probability 1 an amount of \( \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_a + \eta(1-p_a)} \) disadvantaged and high-grade students), we have that \( z_{ud} = \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_a + \eta(1-p_a)} \).

**Advantaged School.** We assume that the advantaged school’s strategies are \( x_{La} = 1; x_{Ha} = 1 \). In the case that the school teaches one high-ability student, the expected payoffs\(^{17}\) are \( \Pi_H^T_a = \mu - c \) for giving extra teaching and \( \Pi_H^T_d = \mu \eta \) for not giving extra teaching. We are assuming that \( \Pi_H^T_a > \Pi_H^N_a \), that is \( \mu - c > \mu \eta \), and therefore \( \mu > \frac{c}{\eta} \). By Assumption 2.1(i), the amount of high-ability students is lower than the labour demand, therefore this strategy occurs \( \lambda p_a \) times, that is for each advantaged and high-ability student who

\(^{16}\)The superscript of the employer’s expected profit indicates the strategy performed by the employer, where \( E \) indicates “to hire” and \( N \) not. The subscript specifies the student’s grade, where \( U \) indicates a high grade and \( D \) a low grade, while \( a \) and \( d \) indicates the student’s social background.

\(^{17}\)The superscript of the (advantaged and disadvantaged) school’s expected profit specifies the school’s strategy, where \( T \) indicates “to provide extra teaching” and \( NT \) not. The subscript indicates the student’s ability and social background.
goes to the job market. Hence \( x_{Ha} = 1 \). In the case that the advantaged school teaches one low-ability student, the expected payoffs are \( \Pi_{La}^T = \mu \eta - c \) for giving extra teaching and \( \Pi_{La}^{NT} = 0 \) for not giving extra teaching. We are assuming that \( \Pi_{La}^T > \Pi_{La}^{NT} \), that is \( \mu \eta - c > 0 \), and therefore \( \mu > \frac{c}{\eta} \). By Assumption 2.(i), the amount of high-grade students is lower than the labour demand, therefore this strategy occurs \( \lambda \eta \) times, that is for each advantaged, high grade and low-ability student who goes to the job market. Hence \( x_{La} = 1 \).

**Disadvantaged School.** We assume that the disadvantaged school’s strategies are \( x_{Ld} = \frac{-\lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)}; x_{Hd} = 1 \). In the case that the school teaches one high-ability student, the expected payoffs are \( \Pi_{Hd}^T = \mu - c \) for giving extra teaching and \( \Pi_{Hd}^{NT} = \mu \eta \) for not giving extra teaching. We are assuming that \( \Pi_{Hd}^T > \Pi_{Hd}^{NT} \), that is \( \mu - c > \mu \eta \), and therefore \( \mu > \frac{c}{1-\eta} \). By Assumption 2.(i), the amount of high-ability students is lower than the labour demand, therefore this strategy occurs \( (1 - \lambda) p_d \) times, that is for each advantaged, high-grade and high-ability student who goes to the job market. Hence \( x_{Hd} = 1 \). The school prefers that a high-ability student would be hired rather than a low-ability student, therefore in turn it consider of providing with extra teaching firstly high-ability students and only residually low-ability students. Thus the proportion of disadvantaged and low-ability students who are considered for receiving extra teaching is given by the labour demand, minus the amount of high-grade advantaged, minus the amount of high-ability disadvantaged, that is \( \frac{-\lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)} \). In the case that the disadvantaged school teaches one low-ability student, the expected payoffs are \( \Pi_{Ld}^T = \mu \eta - c \) for giving extra teaching and \( \Pi_{Ld}^{NT} = 0 \) for not giving extra teaching. We are assuming that \( \Pi_{Ld}^T > \Pi_{Ld}^{NT} \), that is \( \mu \eta - c > 0 \), and therefore \( \mu > \frac{c}{\eta} \). This strategies occurs for an amount of \( \frac{-\lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)} \) disadvantaged and low-ability students, hence \( x_{Ld} = \frac{-\lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)} \). For the remainder of low-ability students, the expected payoffs are \( \Pi_{Ld}^T = -c \) for giving extra teaching and \( \Pi_{Ld}^{NT} = 0 \) for not giving extra teaching. Hence the school never provides extra teaching to these students, as \( -c < 0 \).

**Case 2.** \( p_a > \frac{\eta}{\mu + \eta} \), \( p_a \leq \frac{\eta}{\mu + \eta} \)

**Employer.** We assume that the employer’s strategies are \( z_{Ua} = 1; z_{Da} = 0 \), \( z_{Ud} = \frac{-\lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)} \cdot \mu \eta \), \( z_{Dd} = 0 \). According to the schools’ strategy, in
the case that an advantaged student has a high grade the employer’s beliefs are: \( \pi(\theta_H \mid g_U, a) = \frac{p_a}{p_a + \eta(1-p_a)} \) and \( \pi(\theta_L \mid g_U, a) = \frac{\eta(1-p_a)}{p_a + \eta(1-p_a)} \), while in the case that a disadvantaged student has a high grade, the employer’s beliefs are \( \pi(\theta_H \mid g_U, d) = \frac{p_d}{p_d + \eta_L(1-p_d)} \) and \( \pi(\theta_L \mid g_U, d) = \frac{\eta_L(1-p_d)}{p_d + \eta_L(1-p_d)} \). Thus the expected profits are \( \Pi_{U_a}^E = \frac{p_a}{p_a + \eta(1-p_a)} \nu - \frac{\eta(1-p_a)}{p_a + \eta(1-p_a)} \) for hiring and \( \Pi_{U_a}^N = 0 \) for not hiring one advantaged and high-grade student. On the other hand, the expected profits are \( \Pi_{U_d}^E = \frac{p_d}{p_d + \eta_L(1-p_d)} \nu - \frac{\eta_L(1-p_d)}{p_d + \eta_L(1-p_d)} \nu \) for hiring and \( \Pi_{U_d}^N = 0 \) for not hiring one disadvantaged and high-grade student. Since \( p_a > p_d \), the expected payoff obtained by hiring a high-grade and advantaged student is always higher than the expected payoff obtained by hiring a high-grade and disadvantaged student. Hence \( \Pi_{U_a}^E > \Pi_{U_d}^E \). Also, since we are assuming that \( \Pi_{U_a}^E > \Pi_{U_a}^N \), therefore \( \frac{p_a}{p_a + \eta(1-p_a)} \nu - \frac{\eta(1-p_a)}{p_a + \eta(1-p_a)} > 0 \) and, after few passages, \( p_a > \frac{\eta}{\nu + \eta} \). This strategy occurs for \( \lambda(p_a + \eta(1-p_a)) \) times, that is the number of advantaged and high-grade students. Since, by Assumption 2.(i), this is lower than the labour demand, we have that \( z_{U_a} = 1 \). The remaining part of the labour demand comprehends \( \Phi - \lambda(p_a + \eta(1-p_a)) \) placements, that is there are placements for \( \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \) disadvantaged and high-grade students. For this part of the labour demand, we are assuming that \( \Pi_{U_d}^E = \Pi_{U_d}^N \), therefore \( \frac{p_d}{p_d + \eta_L(1-p_d)} \nu - \frac{\eta_L(1-p_d)}{p_d + \eta_L(1-p_d)} \nu = 0 \) and, after few passages, \( x_{Ld} = \frac{p_d}{(1-p_d) \eta} \). Since this strategy occurs for \( \frac{\Phi - \lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)} \) disadvantaged and high-grade students (i.e., the disadvantaged school provides with extra teaching an amount of \( \frac{\Phi - \lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)} \) disadvantaged and high-grade students), we have that \( x_{Ld} = \frac{\Phi - \lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)} \frac{p_d}{(1-p_d) \eta} \).

**Advantaged School.** For the advantaged school, it matters that \( p_a > \frac{\eta}{\nu + \eta} \). This is the same situation observed in the previous case, thus the advantaged school’s strategy does not change.

**Disadvantaged School.** We assume that the disadvantaged school’s strategies are \( x_{Ld} = \frac{\Phi - \lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d}{p_d + \eta(1-p_d)} \frac{p_d}{(1-p_d) \eta} \nu \), \( x_{Hd} = 1 \). In the case that the school teaches one high-ability student, the expected payoffs are \( \Pi_{Hd}^E = \mu - c \) for giving extra teaching and \( \Pi_{Hd}^N = \mu \eta \) for not giving extra teaching. We are assuming that \( \Pi_{Hd}^E > \Pi_{Hd}^N \), that is \( \mu - c > \mu \eta \), and therefore \( \mu > \frac{c}{1-\eta} \). The amount of high grade and disadvantaged students evaluated by the employer is \( \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \), however, by Assumption 2.(i), the amount of high-ability and advantaged students is lower than the labour demand minus the amount of
advantaged and high-grade students. Therefore this strategy occurs \((1 - \lambda) p_d\) times, hence \(x_{Hd} = 1\). The school prefers that a high-ability student would be hired rather than a low-ability student, therefore in turns it consider of providing with extra teaching firstly high-ability students and only residually low-ability students. Thus the proportion of disadvantaged and low-ability students who are considered for receiving extra teaching is given by the labour demand, minus the amount of high-grade advantaged, minus the amount of high-ability disadvantaged, that is \(\Phi - \lambda(p_a + \eta(1-p_a)) - (1-\lambda)p_d\). In the case that the disadvantaged school teaches one low-ability student, the expected payoffs are 
\[\Pi_{La}^T = z_{Ul} \mu \eta - c\] for giving extra teaching and \(\Pi_{Nd}^{NT} = 0\) for not giving extra teaching. We are assuming that \(\Pi_{La}^T = \Pi_{Nd}^{NT}\), that is \(z_{Ul} \mu \eta - c = 0\), and therefore \(z_{Ul} = \frac{c}{\mu \eta}\). This strategies occurs for an amount of \(\Phi - \lambda(p_a + \eta(1-p_a))\) disadvantaged and high-grade students, hence \(z_{Ud} = \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{\mu \eta}\).

**Case 3.** \(p_a \leq \frac{p}{\nu + \eta}, p_d \leq \frac{p}{\nu + \eta}\)**

**Employer.** We assume that the employer’s strategies are \(z_{Ua} = \frac{c}{\mu \eta}; z_{Da} = 0, z_{Ud} = \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{\mu \eta}; z_{Dd} = 0\). According to the schools’ strategy, in the case that an advantaged student has a high grade the employer’s beliefs are: \(\pi(\theta_H | g_J, a) = \frac{p_a}{p_a + \eta \xi(La)(1-p_a)}\) and \(\pi(\theta_L | g_J, a) = \frac{\eta \xi(La)(1-p_a)}{p_a + \eta \xi(La)(1-p_a)}\), while in the case that a disadvantaged student has a high grade, the employer’s believes are \(\pi(\theta_H | g_J, d) = \frac{p_d}{p_d + \eta \xi(Ld)(1-p_d)}\) and \(\pi(\theta_L | g_J, d) = \frac{\eta \xi(Ld)(1-p_d)}{p_d + \eta \xi(Ld)(1-p_d)}\). Thus the expected profits are 
\[\Pi_{Ua}^E = \frac{p_a}{p_a + \eta \xi(La)(1-p_a)} \nu - 1 - \frac{\eta \xi(La)(1-p_a)}{p_a + \eta \xi(La)(1-p_a)}\] for hiring and \(\Pi_{Ud}^N = 0\) for not hiring one advantaged and high-grade student. On the other hand, the expected profits are 
\[\Pi_{Ud}^E = \frac{p_d}{p_d + \eta \xi(Ld)(1-p_d)} \nu - 1 - \frac{\eta \xi(Ld)(1-p_d)}{p_d + \eta \xi(Ld)(1-p_d)}\] for hiring and \(\Pi_{Ud}^N = 0\) for not hiring one disadvantaged and high-grade student. Since \(p_a > p_d\), the expected payoff obtained by hiring a high-grade and advantaged student is always higher than the expected payoff obtained by hiring a high-grade and disadvantaged student. Hence \(\Pi_{Ua}^E > \Pi_{Ud}^E\). Also, since we are assuming that \(\Pi_{Ua}^E = \Pi_{Ud}^N\), therefore \(\nu - \frac{p_a}{p_a + \eta(1-p_a)} = 0\) and, after few passages, 
\(x_{La} = \frac{p_a}{(1-p_a) \eta}\). This strategy occurs for \(\lambda(p_a + \eta(1-p_a))\) times, that is the number of advantaged and high-grade students. Since, by Assumption 2.(i), this is lower than the labour demand, we have that \(z_{Ua} = 1\). The remaining part of the labour demand comprehends \(\Phi - \lambda(p_a + \eta(1-p_a))\) placements, that is
there are placements for \( \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \) disadvantaged and high-grade students.

For this part of the labour demand, we are assuming that \( \Pi_{Ud}^E = \Pi_{Ud}^N \), therefore \( \frac{p_d}{p_d + \eta x_{La} (1-p_d)} (1-p_d) = 0 \) and, after few passages, \( x_{La} = \frac{p_d}{(1-p_d) \eta} \). Since this strategy occurs for \( \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \) disadvantaged and high-grade students (i.e., the disadvantaged school provides with extra teaching an amount of \( \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \) disadvantaged and high-grade students), we have that \( x_{La} = \frac{\Phi - \lambda(p_a + \eta(1-p_a))}{p_d + \eta(1-p_d)} \).

**Advantaged School.** We assume that the advantaged school’s strategies are \( x_{La} = \frac{p_a}{(1-p_a) \eta} \); \( x_{Ha} = 1 \). In the case that the school teaches one high-ability student, the expected payoffs are \( \Pi_{Ha}^T = \mu - c \) for giving extra teaching and \( \Pi_{Ha}^N = \mu \eta \) for not giving extra teaching. We are assuming that \( \Pi_{Ha}^T > \Pi_{Ha}^N \), that is \( \mu - c > \mu \eta \), and therefore \( \mu > \frac{c}{1-\eta} \). By Assumption 2.(i), the amount of high-ability students is lower than the labour demand, therefore this strategy occurs \( \lambda p_a \) times, that is for each advantaged, high-grade and high-ability student who goes to the job market. Hence \( x_{Ha} = 1 \). In the case that the advantaged school teaches one low-ability student, the expected payoffs are \( \Pi_{La}^T = \mu z_{Ua} \eta - c \) for giving extra teaching and \( \Pi_{La}^N = 0 \) for not giving extra teaching. We are assuming that \( \Pi_{La}^T = \Pi_{La}^N \), that is \( \mu z_{Ua} \eta - c = 0 \), and therefore \( z_{Ua} = \frac{c}{\mu \eta} \). By Assumption 2.(i), the amount of high-grade and disadvantaged students is lower than the labour demand, therefore this strategy occurs \( \lambda \eta (1 - p_a) \) times, that is for each advantaged, high grade and low-ability student who goes to the job market. Hence \( z_{Ua} = \frac{c}{\mu \eta} \).

**Disadvantaged School.** For the disadvantaged school, it matters that \( p_d \leq \frac{\eta}{\nu + \eta} \). This is the same situation observed in the previous case, thus the advantaged school’s strategy does not change.