

IZA DP No. 6805

**Anti-Lemons: School Reputation, Relative Diversity,  
and Educational Quality**

W. Bentley MacLeod  
Miguel Urquiola

August 2012

# Anti-Lemons: School Reputation, Relative Diversity, and Educational Quality

**W. Bentley MacLeod**

*Columbia University,  
NBER and IZA*

**Miguel Urquiola**

*Columbia University  
and NBER*

Discussion Paper No. 6805  
August 2012

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## **ABSTRACT**

### **Anti-Lemons: School Reputation, Relative Diversity, and Educational Quality\***

Friedman (1962) observed that the ability of firms to acquire and maintain reputations for quality is a key ingredient for the efficient provision of goods and services in a market economy. This paper explores the implications of school reputation for skill acquisition and labor market outcomes in an otherwise competitive market. We find that reputation effects can explain several puzzling findings in the economics of education, including the fact that competition can, but does not always, improve skill acquisition. This result follows from an anti-lemons effect (in contrast to Akerlof's lemons effect) that arises when schools can enhance their reputation by positively selecting their students. This leads to excess demand for "high quality" selective schools that drive out non-selective schools. This in turn reduces "relative diversity", a measure of ability dispersion in a school, leading to lower skill acquisition.

JEL Classification: H2, H4, I21, J31

Keywords: education, reputation, competition, labor markets

Corresponding author:

W. Bentley MacLeod  
Department of Economics  
420 West 118th, Mail Code 3308  
New York, NY 10027-7296  
USA  
E-mail: [bentley.macleod@columbia.edu](mailto:bentley.macleod@columbia.edu)

---

\* For discussions and comments we thank David Card, Janet Currie, Dennis Epple, Maria Marta Ferreyra, Roland Fryer, Edward Glaeser, Joseph Hotz, Caroline Hoxby, Larry Katz, Derek Neal, Zvika Neeman, Jonah Rockoff, and Richard Romano. We are particularly grateful to Robert Shimer and the referees for thoughtful comments that significantly improved the paper. We thank our research assistants, Elliot Ash, Uliana Logina, Evan Riehl, Teck Yong Tan and Xing Xia, for excellent work, and gratefully acknowledge the support of the International Growth Center and the Russell Sage Foundation.

*By and large, I'm going to be picking from the law schools that basically are the hardest to get into. They admit the best and the brightest, and they may not teach very well, but you can't make a sow's ear out of a silk purse. If they come in the best and the brightest, they're probably going to leave the best and the brightest, O.K.?* —Antonin Scalia

*"In a way you had more human diversity in the old Harvard," a friend once told me. "It used to be the only thing an incoming class shared was blue blood. But bloodlines are a pretty negligible thing. It allows for an amazing variety in human types. You had real jocks and serious dopes, ..., a few geniuses, and a very high percentage of people with completely average intelligence. But now a majority of kids coming into Harvard share traits that are much more important than blood, race, or class. On a deeper level...they're very much alike. They've all got that same need to be the best, or at least be declared the best by someone in authority."* —Andrew Ferguson in *Crazy U*

## 1. INTRODUCTION

Friedman (1962) observed that the ability of firms to acquire and maintain reputations for quality is a key ingredient for the efficient provision of complex goods and services in a market economy.<sup>1</sup> This paper explores the implications of school reputation for the performance of an education market in which students have different innate abilities and acquire skill as a function of their own effort and the productivity of the school they attend. We find that the introduction of reputation effects into an otherwise competitive market can explain several puzzling empirical results from the literature on the economics of education.

First, the evidence regarding the benefits of school competition is mixed. Some studies suggest significant gains from the introduction of charter schools or vouchers,<sup>2</sup> but others find less promising results.<sup>3</sup> Second, there is evidence that parents and students prefer schools with better peers.<sup>4</sup> Beginning with Coleman (1960), it has been hypothesized that this demand is due to the positive effect that good peers have upon student performance. Yet, recent research suggests that such effects are very small or unstable,<sup>5</sup> and that attending selective schools may offer no net benefit.<sup>6</sup> Third, education performance in major urban areas, particularly for disadvantaged students, is poor relative to less densely populated areas, even though cities such as New York and Chicago have made major investments into school reform.<sup>7</sup>

We find these patterns are consistent with a competitive labor market model in which employers use school reputation to infer student skill, and students of different abilities acquire skills as a function of two factors: the productivity of the schools they attend, and their own effort. If schools cannot select students based upon their innate ability, then a free market tends to raise school productivity. However, if schools

---

<sup>1</sup> See MacLeod (2007) for a review of the literature on reputation and quality assurance.

<sup>2</sup> Angrist, Bettinger, Bloom, Kremer, and King (2002) and Abdulkadiroglu, Angrist, Dynarski, Kane, and Pathak (2011).

<sup>3</sup> Krueger and Zhu (2004) and Hsieh and Urquiola (2006).

<sup>4</sup> Black (1999) and Hastings and Weinstein (2008).

<sup>5</sup> Oreopoulos (2003), Katz, Kling, and Liebman (2006), and Carrell, Sacerdote, and West (2010).

<sup>6</sup> Cullen, Jacob, and Levitt (2006), Abdulkadiroglu, Angrist, and Pathak (2011), and Dobbie and Fryer (2011).

<sup>7</sup> Bowen, Chingos, and McPherson (2009).

use an entrance exam to select students, then competition leads to stratification by ability, reduced student effort, and in some cases lower school productivity.

The results follow from an *anti-lemons* effect that arises when firms can influence their reputation by positively selecting their buyers. This builds upon Akerlof’s (1970) insight that if the quality of goods is difficult to observe, then sellers with high quality goods exit the market, leaving behind only low quality “lemons” for sale. In contrast, the perceived quality of a school depends upon the quality of its *buyers* (as Judge Scalia’s quote above suggests). As a result selective sellers face excess demand and drive out non-selective sellers with worse reputations.

Analogous phenomena are observed in other markets for service goods. For example, restaurants, social clubs, and law firms may be perceived to be of high quality when they serve exclusive clients. What makes education unique is that the industry’s output (student achievement) depends upon both firm (school) productivity and buyer (student) effort—as Bishop (2006) has pointed out, schooling is not sufficient for skill formation; students must also have a desire to learn.

We adapt Holmstrom (1999)’s model<sup>8</sup> to capture the strength of such incentives through a concept we label *relative diversity*: the extent to which students can convey their skill through individual-specific measures of learning, as opposed to school membership. This matters because free entry by schools leads—via the anti-lemons effect—to stratification by ability, and thereby to lower relative diversity and effort.

Such a link between stratification and study time is consistent with circumstantial evidence on student behavior. Hoxby (2009) reports that the selectivity of U.S. colleges has increased over the past decades, while Babcock and Marks (2010) find that during roughly the same period, average hours of study declined.

The anti-lemons effect also implies that parents prefer schools with better reputations, and value better peers *per se*. As a result, the effects of school choice will depend on the details of its implementation. Finally, greater population density facilitates stratification and strengthens the anti-lemons effect, making it harder to improve education in densely populated urban areas.

Many of these implications differ from those implied by the well known Spence signaling model. In that model selection works via the cost of *attending* school. In particular, low ability individuals might refrain from attending a very selective school because of the high personal cost of study at that school. In practice, the more common situation seems to be that these individuals cannot gain admission to such a school.

Our results have various policy implications. For example, the introduction of a standardized graduation exam will raise learning, as suggested by Bishop (2006) and Woessmann (2007). Conversely, the introduction of standardized admissions tests will lead to stratification by ability,<sup>9</sup> and lower student effort. The

---

<sup>8</sup> Previous applications to labor economics include Gibbons and Murphy (1992) and Altonji and Pierret (2001).

<sup>9</sup> As Andrew Ferguson suggests happened with the SAT in the second quote at the start of the paper; see also Lemann (1999).

introduction of competition will more effectively improve performance if it is associated with restrictions on schools' ability to select students, e.g., most states require charter schools to use lotteries if oversubscribed.

Our agenda is as follows. Section 2 sets out a framework in which individuals choose among schools, that in turn make two choices: how much value added to supply, and whether to base their admissions on students' backgrounds and their performance in an admissions test. Only students and schools (and not employers) directly observe these measures of innate ability. Upon graduation, students enter a competitive labor market that pays them its best estimate of their skill. This estimate is based on two signals: individuals' performance on a graduation test, and the reputation of the school they attended, where a school's reputation is simply the expected skill of its graduates.<sup>10</sup> We assume that no individual can single-handedly change her school's reputation. Section 3 introduces student and school preferences and characterizes the efficient allocation. Sections 4 and 5 derive the equilibrium allocations. They begin with a benchmark system in which students are randomly assigned to schools, and show that its performance is suboptimal relative to the efficient allocation. They then analyze whether and how performance improves as one adds elements of choice and competition. Sections 6 and 7 discuss further implications of the model.

## 2. EDUCATION AND WAGES

In this section we set up the basic elements of a framework linking ability and education to wages. We show that a model of school reputation has robust testable implications that do not require solving for the market equilibrium. Rather, we can make predictions regarding student effort and wages conditional upon a given degree of school selectivity.

**2.1. Time line.** The timing of the model is as follows:

- (1) In the first period, student innate ability is realized but not directly observed; it is revealed over time as individuals, schools, and the labor market receive signals that lead them to update their beliefs. Two such signals are an *admissions test* and *family background*, observed by individuals and schools prior to enrollment. These measures are soft information that are not verifiable by employers.
- (2) Schools choose value added and admissions standards, and then students are matched to schools.
- (3) Students make consumption choices and choose study effort. Student skill is realized as a function of student innate ability, student effort, and school value added.
- (4) In the second period, students graduate and enter the labor market. The market observes two signals of their skill: i) the *reputation* of the school each student attended, and ii) an individual-specific

---

<sup>10</sup> More broadly, in a competitive labor market employers have incentives to learn about ability and use all available signals of skill, including socioeconomic background, e.g., Farber and Gibbons (1996), Altonji and Pierret (2001), and Grogger (2011).

measure we term a *graduation test*. The labor market sets wage equal to expected skill given these two signals. Individuals receive their wage and choose their second period consumption.

**2.2. Innate ability and the admissions test.** A large population of students is indexed by  $i$ ; each individual has innate log ability denoted by  $\alpha_i$ , distributed as follows:

$$\alpha_i \sim N\left(0, \frac{1}{\rho}\right),$$

where  $\rho = \frac{1}{\sigma^2}$  is the precision, and  $\sigma^2$  is the variance. As is the norm in labor economics, we will work with log variables and generally omit the term 'log'; we use lowercase terms for logs, and upper case for levels.

Students and schools observe a signal of innate ability we call an *admissions test*. Student  $i$ 's performance on this test is given by:

$$\tau_i = \alpha_i + \epsilon_i^\tau,$$

where  $\epsilon_i^\tau \sim N(0, \frac{1}{\rho^\tau})$  and  $\rho^\tau$  is this measure's precision.

**2.3. Family background.** Let  $y_i$  denote the log family income of student  $i$ . We assume it is normally distributed with mean  $\bar{y}$ . Let  $b_i = y_i - \bar{y}$  denote family *background*—the extent to which family income is above or below the mean. Students and schools observe  $b_i$  (e.g., private schools often use parental questionnaires and interviews), which provides an additional signal of innate ability:

$$\zeta b_i = \alpha_i + \epsilon_i^b,$$

where  $\epsilon_i^b \sim N(0, \frac{1}{\rho^b})$  and  $\rho^b$  is this signal's precision;  $\zeta$  measures the extent to which innate ability varies with background.<sup>11</sup>

Together student  $i$ 's admissions test score and background define her characteristics at the time of admission to school, and are summarized by  $\psi_i = \{\tau_i, b_i\}$ , which has a bi-variate normal distribution with p.d.f. given by  $f(\psi_i)$  (we assume that  $\epsilon_i^\tau$  and  $\epsilon_i^b$  are independently distributed). As stated, we assume that these two traits are observed by students and schools but not by employers.

**2.4. Admissions policies.** The innate ability of the students at a given school depends upon a combination of student choice and school admissions policies. We consider three idealized types of schools characterized by their admissions policies: non-selective schools, exam schools, and elite schools. For simplicity, in this section we consider settings in which school markets consist of only one type of school.

---

<sup>11</sup> In other words  $\zeta$  allows for a general covariance between innate ability and background. The literature has often focused on whether that covariance reflects a causal mechanism (e.g., Solon (1999) and Currie (2009)). For our purposes this is not central—the existence of a correlation means that schools can use background as a signal of ability.

By a *non-selective* school we mean one that accepts all students, regardless of their signals,  $\tau$  or  $b$ . This is in the spirit of public schools' policies in many countries, and is consistent with the use of lotteries among charter schools in the U.S. and voucher schools in Sweden. For now, we further simplify matters by assuming students are randomly assigned to these schools. Hence expected ability at any such school,  $s$ , is:

$$(2.1) \quad \hat{\alpha}_s^{NS} = E\{\alpha_i | s_i = s\} = E\{\alpha_i\} = 0,$$

where the superscript  $NS$  stands for non-selective.

Second, *exam* schools use only the admissions test,  $\tau$ , to select students. For example, many countries feature selective public schools that admit students based only on test performance—background plays no role in their selection mechanisms. We suppose the market is sufficiently thick for exam schools to be completely stratified, namely all the students at school  $s$  have the same admissions score  $\tau_s$ . This simplifies the analysis, allowing us to use a linear updating rule. Specifically, by Bayes' rule the expected innate ability of an individual from an exam school  $s$  with admissions standard  $\tau_s$  is (DeGroot (1972)):

$$(2.2) \quad \begin{aligned} \hat{\alpha}_s^{EX} &= E\{\alpha_i | s_i = s\} \\ &= \frac{\rho^\tau}{\rho^\tau + \rho} \tau_s + \frac{\rho}{\rho^\tau + \rho} E\{\alpha_i\} \\ &= \frac{\rho^\tau}{\rho^\tau + \rho} \tau_s, \end{aligned}$$

where the superscript  $EX$  stands for exam. In (2.2),  $\frac{\rho^\tau}{\rho^\tau + \rho}$  is the weight assigned to the admission test score  $\tau_s$ . Intuitively, the more precise this measure, the greater the weight put on it as opposed to the unconditional mean of ability,  $E\{\alpha_i\} = 0$ . The precision of  $\hat{\alpha}_s^{EX}$  is  $\rho(\alpha|\tau) = \rho^\tau + \rho$ .

Finally, *elite* schools use both the admissions test,  $\tau$ , and the background signal,  $b$ , to select students. One can think of these as resembling private schools, which in many countries can use any criterion (e.g., tests or parental interviews) to select students. This can increase the amount of information conveyed to the market through the admissions process. As before, suppose that the market is sufficiently thick to allow for perfect selectivity—all individuals at school  $s$  have the same ability conditional upon the available information. Hence, for all students that attend an elite school  $s$  the expected ability is:

$$(2.3) \quad \begin{aligned} \hat{\alpha}_s^{EL} &= E\{\alpha_i | s_i = s\} \\ &= \frac{\rho^\tau}{\rho^b + \rho^\tau + \rho} \tau_i + \frac{\rho^b}{\rho^b + \rho^\tau + \rho} \zeta b_i + \frac{\rho}{\rho^b + \rho^\tau + \rho} E\{\alpha_i\} \\ &= \frac{\rho^\tau}{\rho^b + \rho^\tau + \rho} \tau_i + \frac{\rho^b}{\rho^b + \rho^\tau + \rho} \zeta b_i, \end{aligned}$$



where the superscript  $EL$  stands for elite. The weight placed on each signal again depends upon its precision. The precision of  $\hat{\alpha}_s^{EL}$  is the sum of that of the two signals and the prior precision of ability:  $\rho(\alpha|\tau, b) = \rho^b + \rho^\tau + \rho$ .<sup>12</sup> These results are summarized in the following proposition (all proofs are in the appendix).

**Proposition 1.** *Suppose a student with characteristics  $\psi_i$  (admissions test  $\tau_i$  and background  $b_i$ ) attends school in a system featuring only one type of school. Her expected ability conditional upon the type of school she enrolls in satisfies:*

<i>School type:</i>	<i>Non-Selective</i>	<i>Exam</i>	<i>Elite</i>
<i>Expected ability:</i>	$\hat{\alpha}_s^{NS} = 0$	$\hat{\alpha}_s^{EX} = \frac{\rho^\tau}{\rho^\tau + \rho} \tau_s$	$\hat{\alpha}_s^{EL} = \frac{\rho^\tau}{\rho^b + \rho^\tau + \rho} \tau_i + \frac{\rho^b}{\rho^b + \rho^\tau + \rho} \zeta b_i$
<i>Precision of estimate:</i>	$\rho$	$\rho^\tau + \rho$	$\rho^b + \rho^\tau + \rho$

TABLE 1. Expected ability by type of school system attended

**2.5. Skill.** In addition to implementing admissions policies, schools produce educational value added. Specifically, the skill of a student  $i$  who attends school  $s$  depends on: i) her innate ability, ii) her effort, and iii) her school's value added, with a multiplicative form:

$$\Theta_i = A_i E_i V_{s_i}.$$

It is more convenient to carry out the analysis in terms of log variables, accordingly let  $\theta_i$  denote log skill:

$$(2.4) \quad \theta_i = \log(\Theta_i) = \alpha_i + e_i + v_{s_i},$$

where  $\alpha_i = \log(A_i)$  is ability as defined above,  $e_i = \log(E_i)$  is student study effort, and  $v_{s_i} = \log(V_{s_i})$  is the value added that student  $i$  receives at school  $s_i$ .<sup>13</sup> We assume a school provides the same value added,  $v_s$ , to each of its students, and that this level is observed by all actors.

**2.6. Signals of skill: school reputation and the graduation test.** The labor market observes two signals of individual skill. First, employers see the identity of the school attended by each individual and

<sup>12</sup> Notice that when an elite school selects individuals with expected ability  $\hat{\alpha}_s^{EL}$ , it will draw students of various backgrounds, e.g., those with good backgrounds will have lower admissions test scores. Formally, expression (2.3) implies that students at an elite school will lie on a downward sloping line on the  $\tau, \zeta b$  plane. This shows that pure selection can generate a trade-off between test scores and background. Epple and Romano (1998) obtain a similar result. In their case the trade-off emerges from peer effects and price discrimination.

<sup>13</sup> This specification allows for complementarities between ability and value added. Moreover, it is consistent with the empirical literature on the returns to education, where log wages are assumed to be an additively separable function of ability and education (see Mincer (1974) and Card (1999)).

correctly anticipate value added,  $v_s$ , and average effort,  $\hat{e}_s$ , at the school. From Proposition 1, school identity also provides a signal of innate ability,  $\hat{\alpha}_s$ , the expected ability of a randomly selected graduate of school  $s$ . We can thus define a school's *reputation* as the publicly observed expected skill of its graduates:

$$\begin{aligned} R_s &= E\{\theta_i | s_i = s\} = E\{\alpha_i | s_i = s\} + \hat{e}_s + v_s \\ &= \hat{\alpha}_s + \hat{e}_s + v_s. \end{aligned}$$

The assumption of perfect competition in this continuum model implies that individual effort has no effect upon school reputation:

$$(2.5) \quad \frac{\partial R_{s_i}}{\partial e_i} = 0.$$

This reflects that an individual will be unable to affect the average study effort,  $\hat{e}_s$ , at school  $s$ . More generally, schools' reputations depend on the experience of large numbers of graduates/cohorts. Expression (2.5) is a key ingredient because the benefit from attending a school with a good reputation will not stem from any effect upon effort, but rather from the signal regarding innate ability that is transmitted to the market.

Second, the labor market observes an individual-specific measure of skill. We label this measure a *graduation test* and identify it with the statistic  $t_i$ :

$$t_i = \alpha_i + e_i + v_{s_i} + \epsilon_i^t,$$

where  $\epsilon_i^t \sim N(0, \frac{1}{\rho^t})$  is measurement error. Two aspects distinguish this measure from the admissions test,  $\tau$ : i) it is a measure of skill rather than just of innate ability, and ii) it is observed by all agents. As its name indicates, this measure can be motivated using the highly publicized standardized high school graduation exams in countries such as Germany. Other individual-specific measures are seen in other settings and educational levels. For example, in the U.S. college graduates distribute letters of recommendation and lists of honors received, while Economics Ph.D. students distribute "job market papers".

Since the graduation test is an individual-specific measure, it responds to individual effort:

$$(2.6) \quad \frac{\partial t_i}{\partial e_i} = 1.$$

Together with (2.5) this implies that study effort does not respond to rewards associated with school reputation, but does respond to rewards linked to individual-specific measures of performance.

2.7. **Wages.** Upon graduating from school  $s_i$ , an individual with characteristics  $\psi_i$  who attended school  $s$  and obtained a graduation test score  $t_i$  earns a wage  $w(s_i, t_i)$ . Following Jovanovic (1979) we assume a perfectly competitive labor market that sets an individual’s wage equal to the best estimate of her skill given the two signals it observes. Hence we have:<sup>14</sup>

$$(2.7) \quad w(s_i, t_i) = E \{ \theta_i | R_{s_i}, t_i \}.$$

Given that school reputation,  $R_{s_i}$ , and the graduation test,  $t_i$ , are unbiased, normally distributed signals of skill, a person’s wage will be a convex combination of these two measures:

**Proposition 2.** *Suppose the labor market considers an individual  $i$  who attended school in a system featuring only one type of institution. If the student attended a school with reputation  $R_{s_i}$ , and her graduation test score is  $t_i$ , then her wage satisfies:*

<i>School type:</i>	<i>Non-selective</i>	<i>Exam</i>	<i>Elite</i>
<i>Log wage:</i>	$w^{NS}(s_i, t_i) = \pi^{NS} t_i + (1 - \pi^{NS}) R_{s_i}$	$w^{EX}(s_i, t_i) = \pi^{EX} t_i + (1 - \pi^{EX}) R_{s_i}$	$w^{EL}(s_i, t_i) = \pi^{EL} t_i + (1 - \pi^{EL}) R_{s_i}$
<i>Relative diversity:</i>	$\pi^{NS} = \frac{\rho^t}{\rho^t + \rho}$	$\pi^{EX} = \frac{\rho^t}{\rho^\tau + \rho^t + \rho}$	$\pi^{EL} = \frac{\rho^t}{\rho^b + \rho^\tau + \rho^t + \rho}$

TABLE 2. Wage determination given school reputation and the graduation test

Table 2 introduces the notion of *relative diversity*: the precision of the graduation test divided by the precision of the market’s estimate of student skill. Relative diversity is the weight the market places upon the graduation test (as opposed to school reputation) in setting wages. We use the term *relative* because this weight depends upon the precision of the graduation test relative to the precision of the market’s estimate of skill; an increase in the precision of the graduation test, or a decrease in school selectivity can lead to greater relative diversity and increased effort.

Going from left to right, Table 2 illustrates that (keeping the precision of the graduation test fixed) relative diversity is lower in systems in which schools implement more selective admissions. Since students can affect their testing performance but not their schools’ reputation ( $\frac{\partial t_i}{\partial e_i} = 1$  and  $\frac{\partial R_{s_i}}{\partial e_i} = 0$ ), effort increases with relative diversity: the extent to which an individual can affect the market’s impression through the signal she can influence—the graduation test—is greater, and hence so is the incentive to study.

<sup>14</sup> As discussed in footnote 13, we follow the norm in labor economics and carry out the analysis in terms of log wages.

**2.8. Discussion: Education research.** Proposition 2 implies that since wages increase with school reputation, students and parents will value selectivity and better peers *per se*—they will prefer schools with higher achievement even if this advantage does not originate in higher value added or positive peer effects (our setup does not even feature the latter). This result reconciles three important findings in the literature:

- (1) There is clear evidence that parents prefer higher achievement schools. In some cases, this evidence is direct. In the U.S., for instance, it comes from housing valuations (Black (1999)) and parental reactions to information on school performance (Hastings and Weinstein (2008)). In other cases, it can be inferred from reactions to school choice reforms. In Chile, for example, the unrestricted distribution of school vouchers resulted in about 50 percent of all students transferring from the public sector into private schools with higher test scores (McEwan, Urquiola, and Vegas (2008)).
- (2) The evidence on whether higher achievement schools produce higher *value added* is mixed. For example, private schools frequently have higher absolute achievement than public schools, but Neal (2010) and Barrow and Rouse (2009) conclude that this advantage does not systematically extend to value added. This is consistent with the literature studying whether attending a higher achievement school raises testing performance, in which no uniform pattern emerges.<sup>15</sup>
- (3) The evidence on the significance and magnitude of peer effects is mixed. Several studies suggest at most small effects (e.g., Oreopoulos (2003), Katz, Kling, and Liebman (2006)), and recent work finds that the magnitude and very direction of peer effects may be fragile with respect to, for example, classroom composition (Carrell, Sacerdote, and West (2010)).

In short, Proposition 2 shows that a preference for environments with better peers can be rational even if these are not associated with higher value added or learning spillovers.<sup>16</sup>

Proposition 2 also implies that if effort matters for learning, then selectivity—by lowering relative diversity—will decrease skill accumulation. This is consistent with circumstantial and anecdotal evidence on student behavior. In the U.S. the time college students spend studying has declined in the past decades. The dotted line in Figure 2.1 shows this evolution—equivalent to a decline from 40 to 27 hours per week. Hoxby (2009) states that the selectivity of colleges increased over roughly the same period. Figure 2.1 provides one illustration of this by plotting the “market share” of two elite institutions—Princeton and Yale—showing a similar decline. This graph is purely illustrative, but it does show that study time is correlated with college selectivity. Aside from the model presented here, we are not aware of a known explanation for these

---

<sup>15</sup> Specifically, papers like Cullen, Jacob, and Levitt (2005, 2006), Clark (2010), and Abdulkadiroglu, Angrist, and Pathak (2011) find little or no impact. Others, like Pop-Eleches and Urquiola (2011) and Jackson (2010) find positive effects.

<sup>16</sup> This may also explain why even though there is clear evidence that absolute achievement can affect parental school choice, the evidence that school value added does so is weaker (e.g., Rothstein (2006), Mizala and Urquiola (2008), and Deming, Hastings, Kane, and Staiger (2011)).

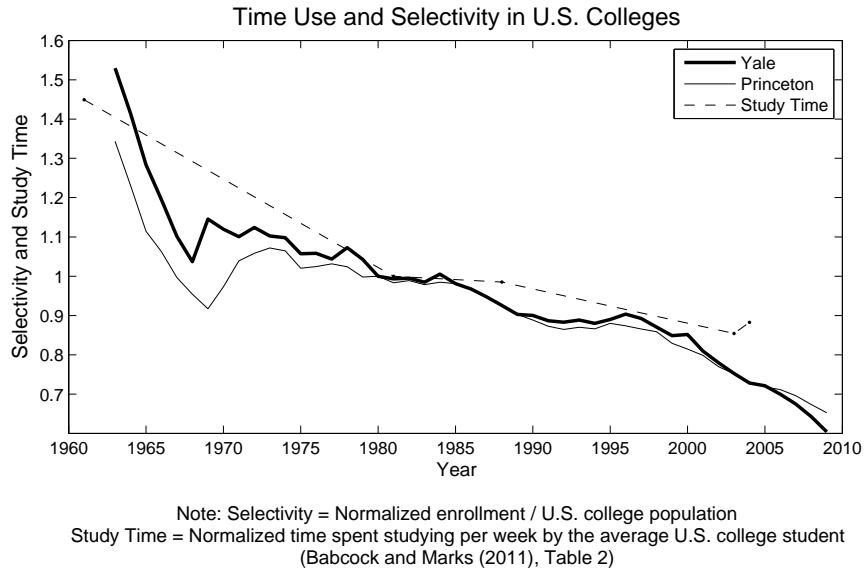


FIGURE 2.1. College Selectivity and Study Time in the U.S.

phenomena.<sup>17</sup> Anecdotally, students in Japan exert high effort to gain admission to elite schools like the University of Tokyo, yet those who succeed are said to lower their study time once there.<sup>18</sup>

This also implies that if school choice reforms are associated with stratification, they may have ambiguous effects on skill, as they may lower study time even if they raise school productivity (we model the latter effect below). Consistent with this, the most extensive implementation of a voucher program, which took place in Chile, had little impact on average achievement even while it resulted in extensive private school entry and a highly stratified school system (Hsieh and Urquiola (2006)).

Proposition 2 implies that introducing precise individual-specific measures of skill raises learning, as emphasized by Bishop (2006). It is consistent with Woessmann (2007)'s finding that at the country level, the existence of standardized graduation or college admissions exams is correlated with better than expected international testing performance. Few observers disagree that such high stakes examinations result in high levels of effort. For instance, countries with such tests often display extensive private tutoring industries that parents use to supplement their children's learning at school (Dang and Rogers (2008)).

<sup>17</sup> Babcock and Marks (2010) report the decline in study time with no explanation. As stated, in Figure 2.1 study time varies from about 40 hours to 27 hours; it is normalized to 1 in 1980. We obtained the data on aggregate college enrollment in the U.S. from the Integrated Post-secondary Education Data System "Fall Enrollment Survey" (IPEDS-EF:86-99) and the Digest of Educational Statistics ([http://nces.ed.gov/programs/digest/d10/tables/dt10\\_198.asp](http://nces.ed.gov/programs/digest/d10/tables/dt10_198.asp)). Princeton and Yale enrollment numbers come from historical publications (details available upon request).

<sup>18</sup> We have not found a documentary source on this; however, our colleague David Weinstein, an expert on Japan, sent us a note stating that dozens of individuals have related to him this perception regarding the University of Tokyo.

**2.9. Discussion: Labor market research.** Proposition 2 shows that *conditional* on years of schooling, school reputation can affect wages. It implies that the identity of the school an individual attends may be a relevant proxy for ability in a wage regression (along with standard controls like experience).<sup>19</sup> While few data-sets contain information on the schools students attended, several influential papers explore situations in which the econometrician observes a measure correlated with ability that is *not* directly observed by the labor market. For example, Farber and Gibbons (1996) know whether a student had a library card when young, and Arcidiacono, Bayer, and Hizmo (2010) know a person’s Armed Forces Qualifying Test (AFQT) score. Because they are not observed by employers, these measures cannot affect wages unless ability is revealed to the market by other means. Proposition 2 implies that an educational system—to the extent it contains selective institutions and graduation assessments—may perform this function.

For example, consider individuals who enter the labor market upon graduation from high school and from college in the U.S. In the case of the former, there are few individual-specific measures of skill ( $t_i$ ), particularly given that the U.S. has no national high school exit exam. Moreover, these individuals tend to come from non-selective schools. In contrast, college graduates come from a sector in which selectivity is much more common, as are individual-specific measures such as letters of recommendation.

These observations suggest that the labor market will have substantial information regarding college graduates’ ability, and little about high school graduates’ (beyond that conveyed by years of schooling). This is consistent with the work by Arcidiacono, Bayer, and Hizmo (2010), which finds that AFQT performance is reflected in wages immediately for college graduates, but only gradually for high school graduates.

A further set of results suggest that these findings at least partially reflect the more selective nature of colleges, as opposed to simply greater availability of information of the type captured by the graduation test. Namely, Saavedra (2009) and Hoekstra (2009) use regression discontinuity designs to show that college selectivity and prestige have a significant and positive effect upon wages early in individuals’ careers. In the first paper the result is for starting wages for college graduates in Colombia; in the second it is for wages 7-12 years after graduation in the U.S.

While our model features only two periods, in practice the labor market gains information about individual ability over time. Thus, the correlation between unobserved measures of ability and wages should increase with experience. This is illustrated by Farber and Gibbon’s (1996) finding that library card ownership when young is uncorrelated with starting wages, but positively correlated with the wages of older workers. Note that the increased availability of information on ability is equivalent to an increase in the precision of  $t$ . This implies that the effect of school reputation on wages should diminish over time. This is consistent with

---

<sup>19</sup> In results available upon request, we formally relate our model to a standard Mincer wage equation. One point that emerges is that variation in selectivity is yet another source of bias in the estimation of the returns to education.

Dale and Krueger’s (2002) finding that school selectivity has little impact upon wages 20 years into workers’ careers.<sup>20</sup> Indeed, the results in Dale and Krueger (2002) are often viewed as showing that school identity does not matter. Rather, these results, along with those on the wage effects early in individuals’ careers by Saavedra (2009) and Hoekstra (2009), are consistent with the Bayesian learning model.

To summarize, a disparate set of studies from the education and labor economics literatures are consistent with the joint hypothesis that wages are set in a competitive market, as in Jovanovic (1979), and that they vary with signals of student skill, such as the reputation of an individual’s alma mater.

### 3. PREFERENCES AND THE EFFICIENT ALLOCATION

This section sets out school and student preferences, and defines a Pareto-optimal allocation. To characterize this efficient allocation, we suppose that a social planner can set student effort and observe all information seen by schools and employers. The result is a benchmark for comparison with later scenarios in which individuals choose effort to maximize private returns, and in which wages are set by employers who only observe school reputation and the graduation test.

Throughout, our goal is to study perfect competition and to explore whether reputation effects alone ensure efficiency. We follow Aumann (1966) and model perfect competition by assuming a continuum of students are matched to a continuum of schools that can perfectly tailor their characteristics to their customers. Our analysis thus complements the work of Arnott and Rowse (1987) and Epple and Romano (1998, 2008) that focuses on peer effects and monopolistic competition between relatively small numbers of schools.<sup>21</sup> In these models there is an efficient scale and hence each school has some market power; our framework assumes perfect competition.

**3.1. Schools.** Schools are uniformly distributed over  $S = (0, \bar{n})$ , where  $\bar{n} > 1$  is the total number of schools serving a population of students normalized to 1.<sup>22</sup> School  $s$  chooses value added per student,  $v_s$ , at per capita cost  $\frac{K(v_s)}{Q_s}$ ;  $Q_s$  is school productivity, which can be high or low:  $Q_s \in \{Q_H, Q_L\}$ , where  $Q_H > Q_L$ . The function  $K(v) > 0$  satisfies  $K'(v)$ ,  $K''(v) > 0$ , and  $\lim_{v \rightarrow -\infty} K(v) = 0$ . Thus, for the same per capita expenditure high productivity schools provide greater value added. The low productivity schools are uniformly distributed over  $S_L = (0, n_L]$ ; the high productivity schools over  $S_H = (n_L, \bar{n} = n_H + n_L)$ . We assume that there is an inelastic supply of schools with  $n_H, n_L > 1$ , so that each type can potentially

<sup>20</sup> Additionally, Lange (2007) finds that approximately 20 percent of the variation in starting wages may be attributed to the signaling effect of schools. Lange emphasizes that this signal is less important for more experienced workers. Finally, Oreopoulos et al. (2006) suggest that the effect of school identity may be heterogeneous. They find that graduates entering into a slack labor market have lower starting wages, but catch up over time. However, for those from the lowest ranked colleges, this catch up is incomplete; for this subgroup there may be a long run effect of college identity.

<sup>21</sup> For example, the empirical section of Epple and Romano (1998) considers 4-10 schools.

<sup>22</sup> This is an open interval because we allocate students based on log ability; thus the support of ability is the open set  $(-\infty, \infty)$ .

supply the whole market. The design of school systems will endogenously determine which schools enter in equilibrium, and whether the high productivity schools dominate the market.

We assume schools operate as publicly financed non-profits. Each receives a per capita payment of  $T\bar{Y}$ , where  $\bar{Y}$  is mean income and  $T \in [0, 1]$  is an income tax levied on each household.<sup>23</sup> The financing follows the student—effectively a Friedman-type voucher system if students are free to choose schools.

Individual  $i$  faces tax liability  $TY_i$ , resulting in a net family log income of  $\log(1 - T) + y_i$ . It is convenient to define the log tax liability by  $m = -\log(1 - T) \simeq T$ , and hence net log income is  $y_i - m \simeq y_i - T$ . Since schools do not make profits, once the tax rate is set value added varies only with school productivity. Formally, given the tax rate  $T$  the value added of school  $s$  with productivity  $Q_s$  solves:

$$(3.1) \quad K(v_s) = T\bar{Y}Q_s.$$

This allows us to write value added as a function of the log tax rate:

$$(3.2) \quad v_s = \bar{K}(T\bar{Y}Q_s) \simeq \bar{K}(m\bar{Y}Q_s),$$

where  $\bar{K} = K^{-1}$ . We assume the approximation is accurate to allow for more elegant expressions.<sup>24</sup>

Finally, since schools are non-profit their pecuniary returns do not vary with the allocation of students. In order to have schools make choices, we must endow them with preferences to break the resulting indifference. We assume that schools value their reputation and enrollment, with preferences given by:<sup>25</sup>

$$(3.3) \quad U_s = \log(n_s \exp(R_s)) = \log(n_s) + R_s.$$

**3.2. Student preferences.** Individuals care about consumption and effort; we follow Epple and Romano (1998) in assuming preferences that take the standard Cobb-Douglas form:

$$U(C^0, C^1, E) = \left(\frac{C^0}{D(E)}\right)^{\gamma_0} (C^1)^{\gamma_1},$$

where  $C^0$  and  $C^1$  are consumption in periods 0 and 1, respectively. Individuals have decreasing marginal utility from income ( $\gamma_0 + \gamma_1 < 1$ ), and  $D(E)$  is the disutility from effort/study that takes place in period 0.

<sup>23</sup> To focus on the effects of reputation, we assume that schools charge a fixed fee and cannot make profits. In MacLeod and Urquiola (2009) we show that for-profit, high productivity schools would charge higher tuition, in some cases capturing the rents rather than supplying higher value added.

<sup>24</sup> All the essential results hold without this simplification.

<sup>25</sup> This assumption reflects that in practice rents accrue to selective schools; for example, high ability alumni might be more able to donate. A preference for high ability students is consistent with some recent research, e.g., Hatfield, Kojima, and Narita (2011). Turner (2012) documents that higher educational institutions extend greater financial aid to more able students conditional on their Pell grant status. Pop-Eleches and Urquiola (forthcoming) document that in Romania more senior teachers gravitate towards more able students.



We suppose, as in the labor literature, that individuals maximize the expected value of log utility:

$$(3.4) \quad u(c^0, c^1, e) = \gamma_0(c^0 - d(e)) + \gamma_1 c^1,$$

where  $c^t = \log(C^t)$ ,  $t = 0, 1$ ;  $d(e) = \log(D(\exp(e)))$ , and  $e = \log(E)$ . We assume  $\lim_{e \rightarrow -\infty} d(e) = -\infty$ ,  $d'(e) > 0$ , and  $d''(e) > 0$ . We abstract from savings and borrowing, and hence individuals face period-specific budget constraints:

$$(3.5) \quad c_i^0 \leq y_i - m \quad \text{and} \quad c_i^1 \leq w_i.$$

Thus, the relevant choices for an individual are effort,  $e$ , and the tax rate,  $m$ , which in turn determines value added as a function of school productivity.

**3.3. Allocations.** An *allocation* specifies the tax rate and the distribution across schools of students and student effort. We denote it by  $\Gamma = \{m, e_{\psi s}, f(\psi|s)\}_{\psi \in \Psi, s \in S}$ , where  $e_{\psi s}$  is the effort chosen by students with characteristics  $\psi$  at school  $s$ , and  $f(\psi|s)$  describes the distribution of students at school  $s$ . Let  $n_s = \int_{\psi \in \Psi} f(\psi|s) d\psi \leq 1$  denote the number (density) of students at school  $s$ . A feasible allocation is one that satisfies the budget constraint for schools (3.2) and in which the match of students to schools satisfies:

$$(3.6) \quad f(\psi) = \int_{s \in S} f(\psi|s) ds, \quad n_s \leq 1, \quad \text{and} \quad \int_0^{\bar{n}} n_s ds = 1.$$

The first condition requires that the allocation be consistent with the aggregate distribution of types. The second reflects the school capacity constraint. The third condition requires all students to attend school. An allocation implies a unique payoff for each student, in which the budget constraints (3.5) hold with equality. Consumption is  $c_{\psi s}^0 = y_i - m$  in period 0.

The social planner observes all information available to schools and employers. Thus consumption in period 1 equals the wage that would prevail if all schools were “elite” or perfectly selective (Section 2):<sup>26</sup>

$$(3.7) \quad \begin{aligned} c^1(\psi, t, s) &= E\{\alpha|\psi, t\} + e_{\psi s} + v_s \\ &= \frac{\rho^\tau}{\rho^t + \rho^b + \rho^\tau + \rho} \tau + \frac{\rho^b}{\rho^t + \rho^b + \rho^\tau + \rho} \zeta b + \frac{\rho^t}{\rho^t + \rho^b + \rho^\tau + \rho} (t - e_{\psi s} - v_s) + e_{\psi s} + v_s. \end{aligned}$$

We evaluate the efficiency of allocations after student characteristics are determined and at the time students are assigned to schools.<sup>27</sup> At this point individuals do not know their future graduation test

<sup>26</sup> If schools are elite, reputation reveals all the information schools gather as part of admissions (Section 2).

<sup>27</sup> See Holmström and Myerson (1983) for a discussion of efficiency as function of different amounts of information.

realization,  $t_i$ . Hence their expected period 1 consumption is:

$$(3.8) \quad c_{\psi s}^1 = \hat{\alpha}_\psi + e_{\psi s} + v_s,$$

where  $\hat{\alpha}_\psi = E\{\alpha|\psi\}$  is the expected innate ability given student characteristics (Section 2).

Thus, given a feasible allocation  $\Gamma$ , upon being admitted to school  $s$  a student's payoff is:

$$(3.9) \quad u(\psi, s|\Gamma) = \gamma_0 (y_i - m - d(e_{\psi s})) + \gamma_1 (\hat{\alpha}_\psi + e_{\psi s} + v_s).$$

The expected utility of a student with characteristics  $\psi$  is thus:

$$u(\psi|\Gamma) = \int_{s \in S} u(\psi, s|\Gamma) f(\psi|s) ds.$$

Given this payoff, an allocation  $\Gamma$  is *Pareto efficient* if there does not exist another feasible allocation  $\Gamma'$  such that  $u(\psi|\Gamma') \geq u(\psi|\Gamma)$  for all  $\psi \in \mathfrak{R}^2$ , with a strict inequality for a positive measure of  $\psi$ .

**Proposition 3.** *A feasible allocation  $\Gamma$  is Pareto efficient if and only if (up to a set of measure zero):*

1. *The marginal cost of effort is set equal to the rate of time preference:  $e_{\psi s} = e^*$  where  $e^*$  satisfies:*

$$d'(e^*) = \frac{\gamma_1}{\gamma_0}.$$

2. *Students are allocated to high productivity schools:  $n_s = 0$  for  $s \in S_L$ .*
3. *The optimal tax rate  $m^H$  is independent of student characteristics and satisfies:*

$$\bar{K}'(m^H Q_H \bar{Y}) = \frac{\gamma_0}{\gamma_1} \times \frac{1}{Q_H \bar{Y}}.$$

The first condition is that the marginal cost of study equals the future return—implying that effort is independent of background and ability. The second requires that only high productivity schools be utilized. Finally, taxes are set so that the future marginal return is equal to the marginal cost. The optimal tax rate depends on school productivity, but not on individual characteristics. Given that  $d'' > 0$  and  $\bar{K}'' < 0$  it follows that when students are more patient ( $\frac{\gamma_1}{\gamma_0}$  is higher), they work harder and enjoy higher value added at an efficient allocation. Similarly, greater school productivity or higher average family income result in higher value added, with no effect on study effort.

#### 4. SCHOOL CHOICE WITH NON-SELECTIVE SCHOOLS

Would competition achieve the efficient allocation? This section begins to address this question by studying two scenarios. First, a “public” system without choice: students are randomly assigned to schools and

school productivity is randomly drawn. Second, a system in which students can request specific non-selective schools and slots are allocated via lottery if there is excess demand. This approximates policy in U.S. charter and Swedish voucher schools (Böhlmark and Lindahl (2012)). Together, these scenarios show that choice can raise school productivity and expenditure on value added. As a preliminary, Section 4.1 defines incentive compatible allocations and student preferences as a function of school reputation and relative diversity.

**4.1. Incentive compatible allocations.** Since students care only about effort and consumption (3.4) characterizing their payoffs given an allocation requires determining: i) the incentives for effort they face, and ii) the wages they can expect upon graduation. Consider a student  $\psi = \{\tau, b\}$  who attends school  $s$  under allocation  $\Gamma = \{m, e_{\psi s}, f(\psi|s)\}_{\psi \in \Psi, s \in S}$ . Let  $w_{\psi s}$  be her expected wage given this allocation. Section 2 showed how a school's relative diversity,  $\pi_s$ , affects students' marginal incentives in a log normal framework. Given our assumption that the labor market observes only school identity and the graduation test, and given the linear relationship between skill and the graduation test, we have that  $\frac{\partial E\{\alpha|t, s\}}{\partial t}$  is independent of  $\psi$ . Thus we can define relative diversity as the marginal effect of effort upon wages:<sup>28</sup>

$$(4.1) \quad \pi_s = \frac{\partial w_{\psi s}}{\partial e_{\psi s}} = E \left\{ \frac{\partial E\{\alpha|t, s\}}{\partial t} \Big| \psi, s \right\}.$$

This motivates the following definition:

**Definition 1.** An allocation  $\Gamma$  is *incentive compatible* only if:

$$(4.2) \quad d'(e_{\psi s}) = \pi_s \frac{\gamma_1}{\gamma_0}.$$

Given a distribution  $f(\psi|s)$  and tax rate  $m$ , effort is determined by (4.2). Note that effort at school  $s$  is independent of individual characteristics. Hence we drop dependence of effort upon  $\psi$  without loss of generality and denote an incentive compatible allocation by  $\Gamma = \{m, f(\psi|s)\}_{\psi \in \Psi, s \in S}$ .

Checking if an incentive compatible allocation is an equilibrium requires solving for the payoffs a student with characteristics  $\psi$  obtains when allocated to school  $s$ . Consider first schools for which  $n_s > 0$ ; reputation and relative diversity are well defined at these. Since a single student cannot affect these school traits, they may be taken as given by the allocation. School reputation is thus:

$$R_s = \hat{\alpha}_s + \hat{e}_s + v_s,$$

---

<sup>28</sup> Proposition 2 works out this parameter for the cases considered below. In these, the effect of  $t_i$  is linear:  $\partial^2 E\{\alpha_i|s, t_i\} / \partial t_i^2 = 0$ , such that the marginal effect of study time is independent of student characteristics. Note also that relative diversity is well defined as long as there are some students attending a given school.

where  $\hat{e}_s$  is the expected effort for students at school  $s$ . For student  $\psi$ , the expected graduation test score is:

$$E\{t|\psi\} = \hat{\alpha}_\psi + e_{\psi s} + v_s,$$

where  $e_{\psi s}$  is effort chosen by the student  $\psi$  as school  $s$ . The utility of a student  $\psi$  who chooses effort  $e_{\psi s}$  at school  $s$  under allocation  $\Gamma$  is:

$$(4.3) \quad u(\psi, s, e_{\psi s}) = \gamma_0 \{y_i - m - d(e_{\psi s})\} + \gamma_1 \{\pi_s (\hat{\alpha}_\psi + e_{\psi s} + v_s) + (1 - \pi_s) R_s\}.$$

It follows that optimal effort  $e_{\psi s}$  satisfies the incentive compatibility condition, and hence expected school effort solves:  $\hat{e}_s = e_{\psi s}$ . Expression (4.3) has the following implications regarding student preferences:

- (1) Students prefer higher value added:

$$\frac{\partial u}{\partial v_s} = \gamma_1.$$

- (2) Students prefer more able peers and schools with better reputations:

$$\frac{\partial u}{\partial R_s} = \frac{\partial u}{\partial \hat{\alpha}_s} = \gamma_1 (1 - \pi_s).$$

- (3) The preference for relative diversity is ambiguous. By the envelope theorem small changes in relative diversity have little impact upon effort at an incentive compatible allocation, hence we have:

$$\frac{\partial u}{\partial \pi_s} = \hat{\alpha}_\psi - \hat{\alpha}_s,$$

i.e., a student prefers higher relative diversity *if and only if* her expected ability is higher than the mean at her school.

When Friedman (1962) advocated for competition, he had in mind that reputation reflects value added, as in point (1). The fact that student ability is imperfectly measured implies that reputation also reflects student composition, and thus students also prefer more able peers. Finally, students with ability above their school average prefer more “diverse” schools. Similarly, a more precise graduation test increases relative diversity, which is preferred by more able students, but disliked by the less able. This may explain the controversy surrounding the recent move by Princeton University to increase the variance in grades.<sup>29</sup>

Finally, for schools with no students ( $n_s = 0$ ) Bayes theorem provides no guidance on how to set reputation and relative diversity. This issues arises in all market with reputation effects—in the absence of past

---

<sup>29</sup> Specifically, some students complained that a policy to reduce the number of A grades awarded hurt their employment and graduate school prospects. Some implied they would rather be evaluated on the basis of their enrollment at Princeton than on the basis of a grade point average with increased variation (see *Type-A-Plus Students Chafe at Grade Deflation* in the New York Times, January 29, 2010).

experience the market cannot rationally evaluate new entrants. For each case we consider we will have to define appropriate off-equilibrium beliefs that support the equilibrium.

**4.2. A public system.** Consider now a “Public school” (PS) system without choice—at step 2 in the timeline (Section 2) students are randomly allocated to schools in the set  $S$ . This implies that each school has the same: size,  $n_s = \frac{1}{\bar{n}}$ ; distribution of students,  $f^{PS}(\psi|s) = \frac{f(\psi)}{\bar{n}}$ ; average ability,  $\hat{\alpha}_s = 0$ ; and relative diversity,  $\pi^{NS} = \frac{\rho^t}{\rho + \rho^t}$  (Section 2). Hence effort at all schools satisfies:

$$(4.4) \quad d'(e^{NS}) = \pi^{NS} \frac{\gamma_1}{\gamma_0}.$$

Since  $\pi^{NS} < 1$ , study time is lower than at the efficient allocation.

We assume school productivity is randomly drawn (e.g., the government might be unable to exclude low productivity producers). The tax is set by the median voter, such that its level depends upon whether most schools are of high or low productivity. To make things stark we suppose that  $n_L > n_H$  and  $\frac{n_L}{n_L + n_H} > \frac{1}{2}$ . Hence the median voter attends a low productivity school and sets the tax to solve:

$$\bar{K}'(m^L Q_L \bar{Y}) = \frac{\gamma_0}{\gamma_1} \times \frac{1}{Q_L \bar{Y}}.$$

Thus value added at each type of school is given by

$$\begin{aligned} v_s &= v^{L*} = \bar{K}(m^L Q_L \bar{Y}), \quad \forall s \in (0, n_L] \\ v_s &= v^{LH} = \bar{K}(m^L Q_H \bar{Y}), \quad \forall s \in (n_L, \bar{n}), \end{aligned}$$

i.e., value added is lower than at the first best for students fortunate to be in high productivity schools.

To summarize, let  $\Gamma^{PS} = \left\{ m^L, f^{PS}(\psi|s)_{\psi \in \Psi, s \in S} \right\}$  denote the allocation under a public system (PS). Relative to the first best, this setup has lower effort, financing, and average school productivity. Could school choice improve matters?

**4.3. Non-selective schools with choice.** Now consider a system composed of *non-selective* schools that allows for student choice—we label it PC for “Public choice”. Students can apply to any school they desire, with oversubscribed slots allocated via lottery. If a school has no students, we suppose its relative diversity is  $\pi^{NS}$  and its reputation is  $R_s = 0$ . These beliefs are consistent with the market’s expectation that there is no selection in the allocation of students to schools, as is necessary to allow new, higher productivity schools, to compete with incumbents. Given this, a public choice equilibrium is defined as follows.

**Definition 2.** A feasible, incentive compatible allocation  $\Gamma^{PC} = \{m^{PC}, f^{PC}(\psi|s)_{\psi \in \Psi, s \in S}\}$  with associated individual payoffs  $u(\psi, s)$  is a *public choice equilibrium* if:

- (1) A student cannot be made better off by choosing a school with excess capacity: let  $S^0 = \{s | n_s < 1\}$  be the set of schools with empty slots; then for every  $\psi \in \Psi$  and  $s \in S$  such that  $f^{PC}(\psi|s) > 0$  :  $u(\psi, s) \geq u^0(\psi) \equiv \sup_{s' \in S^0} u(\psi, s')$ .
- (2) All desirable, oversubscribed schools randomly select students: for each  $s \in S^1 = \{s | n_s = 1\}$  there is a  $\beta_s > 0$  such that  $f^{PC}(\psi|s) = \beta_s f(\psi)$  for every  $\psi \in \Psi$  satisfying  $u(\psi, s) \geq u^0(\psi)$ .
- (3) The median voter selects the tax rate  $m^{PC}$ .

At over-subscribed schools  $n_s = 1$ , mean ability is zero and relative diversity is  $\pi^{NS}$ . Since students can choose any school, it must be the case that all schools with  $n_s \in (0, 1)$  yield the same payoff. From (4.3), if two schools have different relative diversity, one will be oversubscribed, from which it follows—given lottery admissions—that relative diversity and reputation must be the same at all schools. Additionally, given that there is an excess supply of high productivity schools, the only possible equilibrium is for all schools to be of high productivity and have the same reputation and relative diversity.<sup>30</sup> Since the median voter attends a high productivity school, taxes are set at:  $m^{PC} = m^H$ . This yields the following proposition:

**Proposition 4.** *The Public school (PS) system is Pareto dominated by a Public choice (PC) equilibrium characterized by feasible and incentive compatible allocations satisfying:*

- (1) *Low productivity schools have no students:  $n_s = 0$  for  $s \in S_L$ .*
- (2) *For high productivity schools  $s \in S_H$ ,  $f^{PC}(\psi|s) = n_s f(\psi)$ .*
- (3) *Taxes are set at the efficient level for high productivity schools:  $m^{PC} = m^H$ , such that value added at all schools with students is  $v^{H*} = \bar{K}(m^H Q_H \bar{Y})$ .*
- (4) *Relative diversity at all schools with students is  $\pi^{NS}$ , and hence study time is  $e^{PC} = e^{NS}$ .*

Thus choice can improve school productivity. It does not achieve efficient effort because the market places the weight  $\pi^{NS} < 1$  upon the graduation test, resulting in an incentive to study  $e^{NS} < e^*$ .

We have modeled the supply of schools in stark terms to highlight some ingredients that are necessary for school choice to raise performance. Note that if parents are poorly informed (e.g., if they are unable to identify high productivity schools) then choice might not raise performance. Second, given that students value peer quality, successful entry will depend upon expectations regarding who will attend a given school. For example, some charter schools rely on rules that ensure admission for the children of founding parents

<sup>30</sup> If high productivity schools were in fixed supply and unable to cover the market, and further allowed to set tuition, their owners would accrue a rent; see Epple and Romano (1998) and MacLeod and Urquiola (2009).

to help fix initial expectations regarding student composition. In other cases, in the absence of past history, successful entry depends upon the ability of entrants to manipulate the beliefs of parents rather than upon a proven history of high performance.

## 5. SCHOOL CHOICE WITH SELECTIVITY: THE ANTI-LEMONS EFFECT

Section 4 explored settings in which schools are non-selective. We now consider a “Free entry” (FE) system in which, for simplicity, schools can choose to be non-selective or to operate as “elite” schools that exploit all available information in admitting students (Section 2).

**5.1. Optimal admissions standards.** Before defining the strategies of elite schools, we note a key feature of selective systems in our setup—individuals prefer to attend more rather than less selective schools. Specifically, consider an elite school  $s$  that admits a range of students: those whose expected ability  $\hat{\alpha}_i$  is in the interval  $[\underline{\alpha}_s, \bar{\alpha}_s]$ , with a continuous distribution over this range given by  $h_s(\hat{\alpha})$ , with  $h_s(\underline{\alpha}_s) > 0$ . We suppose that competition by more selective schools prevents the school from raising  $\bar{\alpha}_s$ . However, the school is free to become more selective by increasing  $\underline{\alpha}_s$ . Student utility at this elite school is given by:

$$(5.1) \quad u(\psi, e_{\psi s}, s) = \gamma_0 \{y_{\psi} - m - d(e_{\psi s})\} + \gamma_1 \{\pi_s (\hat{\alpha}_{\psi} + e_{\psi s} + v_s) + (1 - \pi_s) R_s\}.$$

By the envelope theorem, we can ignore the effect of a small increase in selectivity upon effort; similarly, the effect on relative diversity is small since  $h_s(\cdot)$  has a continuous density. However, increasing selectivity via  $\underline{\alpha}_s$  has a first order positive effect upon reputation, and hence a school that enters with a slightly higher minimum score would attract all students above this bound. Thus, we have the following proposition.

**Proposition 5.** *Suppose that school  $s$  admits only students with expected innate ability  $\hat{\alpha}_i \in [\underline{\alpha}_s, \bar{\alpha}_s]$ , with a continuous distribution over this interval given by  $h_s(\hat{\alpha})$ , with  $h_s(\underline{\alpha}) > 0$ ; then increasing selectivity,  $\underline{\alpha}_s$ , raises the payoff of all students remaining in the school.*

Hence, in equilibrium elite schools will choose to be perfectly selective, as we assumed in Section 2. In other words, while schools always prefer a better reputation, given competition they will sort and specialize in a specific student type, as in the Gabszewicz and Thisse (1979) model of vertical differentiation. Note that this prediction is consistent with Figure 2.1. As the college population has grown there has not been a corresponding increase in the size of two elite schools—Princeton and Yale.<sup>31</sup>

---

<sup>31</sup> Bowen, Chingos, and McPherson (2009) show that this holds for major universities, including state institutions.

In light of Proposition 5 we assume each school chooses an admission criteria  $ac_s \in \{\mathfrak{R}, NS\}$ . If  $ac_s = NS$  the school is non-selective and uses a lottery if oversubscribed. If  $ac_s \in \mathfrak{R}$  then it operates as an elite school and admits only students with expected ability  $ac_s = \hat{\alpha}_s$ . Formally, the matching process is as follows:

- (1) Schools  $s \in S$  decides whether to enter and then selects admissions criteria  $ac_s$ , correctly anticipating the student response.
- (2) Students choose the set of schools they prefer from those for which they meet the admissions criteria.
- (3) Students are matched, with a lottery used at over subscribed schools. Students who are not admitted to a school of their choice are randomly allocated to non-selective schools with excess capacity.

By construction, if school  $s$  chooses admissions criteria  $ac_s = \hat{\alpha}_s$  and enrolls students, then its reputation is  $R_s = \hat{\alpha}_s + \hat{e}_s + v_s$ , and its relative diversity is  $\pi_s = \pi^{EL}$  (Proposition 2). We suppose that this also defines the market beliefs when there are no students at a school.

If a school sets strategy  $ac_s = NS$  and has no students in equilibrium, then we assume that its reputation is the same as that of the non-selective schools with students, adjusting for differences in value added. If there are no non-selective schools in equilibrium, then we assume that  $R_s = -\infty$ , i.e., only the worst students would attend non-selective schools. However, this does not preclude entry. If there is insufficient capacity then excess students would be allocated to new entrants. Given that schools prefer some students rather than none (3.3), anticipating this would result in the entry of some schools.

**5.2. Equilibrium with only selective schools.** In this setting there always exists an equilibrium featuring only elite schools. To construct this allocation define the density function for expected ability:

$$(5.2) \quad g(\hat{\alpha}) = \int_{\psi \in \{\psi | E\{\alpha | \psi\} = \hat{\alpha}\}} f(\psi) d\psi,$$

and let  $G(\hat{\alpha})$  be the associated cumulative distribution function. Given that there is an excess supply of high productivity schools, there cannot be an equilibrium featuring low productivity schools.

Let  $0 < n_s \leq 1$  for  $s \in S_H$  be any continuously differentiable function such that  $\int_{s \in S_H} n_s ds = 1$ . This defines the number of students allocated to school  $s$ . We now match students to schools as follows. Let  $\hat{\alpha}(s)$  be the admissions standard for school  $s$ , and choose  $s_0$  such that  $\int_{-\infty}^{s_0} n_s ds = \frac{1}{2}$ . Set  $\hat{\alpha}(s_0) = 0$ —this is the ability of the median student. For  $s \in (n_L, \bar{n})$  let  $\hat{\alpha}(s)$  be the solution to:

$$\frac{d\hat{\alpha}(s)}{ds} = \frac{n_s}{g(\hat{\alpha}(s))}$$



with the boundary condition  $\hat{\alpha}(s_0) = 0$ .<sup>32</sup> This matches students to schools so that school  $s$  has entrance requirement  $\hat{\alpha}(s)$ , with  $n_s \leq 1$  students.

We can see that this is an equilibrium as follows. First, observe that for a school in  $s \in (0, n_L]$  to successfully enter it must offer an admissions policy and value added that make students better off than at the existing equilibrium. If it entered with an admissions policy of  $\hat{\alpha}_s$ , it could not make any type strictly better off since all are matched to high productivity schools. Under the belief that only the worst students attend non-selective institutions, a school cannot successfully enter with a non-selective strategy.

Let  $f^{FE}(\psi|s, n_s)$  be the assignment of students to schools we have just defined. Given that all students attend high productivity schools, the median voter selects the high tax rate  $m^H$ . Given that all schools are perfectly selective, the relative diversity at each school is  $\pi^{EL}$  (Proposition 2), and effort is  $e_s = e^{EL}$ , where  $e^{EL}$  is the corresponding solution to the incentive constraint (4.2). Thus:

**Proposition 6.** *There exists a free entry (FE) equilibrium featuring only elite high productivity schools with  $n_s$  students at each school, as given by the incentive compatible allocation:  $\Gamma^{FE} = \{m^H, f^{FE}(\psi|s, n_s)\}_{\psi \in \Psi, s \in S}$ .*

The fact that schools target particular strata of students means that any competition is within a given segment. Note that the demand for schools that target high ability students here is not driven by value added, but by reputation that rests upon selectivity. This result helps explain the stability of school rankings.<sup>33</sup> Once a school is perceived to be selective, students will continue to rank it highly because of the expectation that able students attend the school. This behavior persists even if value added does not vary across schools.

Proposition 6 offers a counterpart to Akerlof’s *lemons* effect. In his model, sellers stay out of the market due to the presence of low quality sellers. In our case, the quality of the good is determined by the quality of the *buyers*—selective sellers face excess demand from high quality buyers, leading to further entry by selective sellers. In this case, it is the “low quality” sellers that are driven from the market, and thus we call this an *anti-lemons* effect.

**5.3. A hybrid equilibrium.** In Akerlof’s model an equilibrium can feature some trade as long as the high quality sellers are sufficiently motivated to sell. We have an analogous case in the anti-lemons setup when there is some cost of attending an elite school. This allows for a hybrid equilibrium featuring both selective and non-selective schools, a common outcome in education markets. For example, we could assume that elite schools are private and charge tuition, whereas non-selectives are funded only via the tax (a realistic setup

<sup>32</sup> The number of students with ability greater than or equal to  $\hat{\alpha}(s)$  is  $1 - G(\hat{\alpha}(s))$ . We ensure there are  $n_s$  students at school  $s$  by requiring  $\frac{dG(\hat{\alpha}(s))}{ds} = n_s$ . The existence and uniqueness of a solution require the standard conditions for the solution of an ordinary differential equation (Hirsch and Smale (1974)).

<sup>33</sup> This may also explain why the perceived quality of universities is correlated with their age.

in many countries.) In order to explore the effect of population density, we suppose that there are travel costs associated with attending elite schools. For example, there might be a non-selective school in each neighborhood, whereas elite schools might draw from a larger catchment area. Areas with lower population density will have higher transportation costs associated with elite school attendance.

To build the hybrid equilibrium, we begin by noting that the most able students always wish to attend elite schools. Hence if an equilibrium exists it will be characterized by a cutoff  $\alpha^{EL}$ : students with expected innate ability above this ( $\hat{\alpha} \geq \alpha^{EL}$ ) attend a selective school, and those below attend a non-selective school. By the previous arguments, only high productivity schools will enter the market, and the tax rate will be  $m^H$ . Relative diversity at the elite schools is  $\pi^{EL}$  (Proposition 2), and hence incentive compatible effort is  $e^{EL}$  as in the previous section.

In the non-selective sector the distribution of students is given by a normal distribution truncated at  $\alpha^{EL}$ . The expected ability of an individual in this sector is:

$$Z(\alpha^{EL}) = E\{\alpha | \hat{\alpha} \leq \alpha^{EL}\}.$$

In the appendix we discuss the properties of  $Z$ , and show for  $\alpha^{EL} < 0$ :

$$(5.3) \quad \alpha^{EL} > Z(\alpha^{EL}) > \frac{cov(\alpha, \hat{\alpha})}{\alpha^{EL}} + \alpha^{EL}.$$

This implies that controlling for value added and effort, the expected skill of a student from the non-selective sector is always lower than that of the weakest student at an elite school.

The first order condition for effort in the non-selective sector is:

$$(5.4) \quad d'(e^{NS^*}(\alpha^{EL})) = \frac{\gamma_1}{\gamma_0} \pi^{NS^*}(\alpha^{EL}),$$

where  $NS^*$  denotes this is a hybrid market. Relative diversity is given by:

$$\pi^{NS^*}(\alpha^{EL}) = E\left\{\frac{\partial w^{NS^*}(\alpha^{EL}, t_i)}{\partial t_i} | \psi_i\right\}.$$

The term in brackets is the effect of the graduation test upon wages. Under the assumptions we have made, this derivative exists, is a continuous function of the cutoff  $\alpha^{EL}$ , and is independent of  $\psi_i$ . The following proposition describes the qualitative properties of relative diversity at the non-selective schools:

**Proposition 7.** *Consider students in non-selective schools when these co-exist with elite schools as defined by a cutoff  $\alpha^{EL}$ . Relative diversity,  $\pi^{NS*}(\alpha^{EL})$ , is increasing in  $\alpha^{EL}$  and satisfies:*

$$\frac{\rho^t}{\rho + \rho^t} > \pi^{NS*}(\alpha^{EL}) > \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t},$$

$$\lim_{\alpha^{EL} \rightarrow \infty} \pi^{NS*}(\alpha^{EL}) = \frac{\rho^t}{\rho + \rho^t}; \quad \text{and} \quad \lim_{\alpha^{EL} \rightarrow -\infty} \pi^{NS*}(\alpha^{EL}) = \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t}.$$

Thus growth of the selective sector has a negative externality on the non-selective sector.

The final step is to provide conditions under which a hybrid equilibrium exists. Let the log transportation cost of using an elite school be given by  $tc^{EL}$ . For simplicity we assume that transportation costs reduce school value added. For example, time otherwise devoted to instruction might have to be spent traveling to school. Formally, we let value added be defined by:

$$v_s = \begin{cases} \bar{K}(mQ_s\bar{Y}), & \text{if } ac_s = NS, \\ \bar{K}(mQ_s\bar{Y}) - tc^{EL}, & \text{if } ac_s \in \mathfrak{R}. \end{cases}$$

Given this, Proposition 8 provides necessary and sufficient conditions for the existence of a hybrid equilibrium:

**Proposition 8.** *There exists a hybrid equilibrium featuring (high productivity) elite and non-selective schools, a tax rate  $m^H$ , and an ability cutoff level  $\alpha^{EL}$ , if and only if the reputational gain from attending the marginal elite school is exactly offset by the lower effort and transportation costs. Formally,  $\alpha^{EL}$  satisfies:*

$$(5.5) \quad (1 - \pi^{NS}(\alpha^{EL})) (\alpha^{EL} - Z(\alpha^{EL})) = g(\alpha^{EL}) + tc^{EL},$$

where

$$g(\alpha^{EL}) = e^{NS*}(\alpha^{EL}) - d(e^{NS*}(\alpha^{EL})) \frac{\gamma_0}{\gamma_1} - \left( e^{EL} - d(e^{EL}) \frac{\gamma_0}{\gamma_1} \right)$$

is the net gain in effort from attending the non-selective sector.

The left hand side of (5.5) is the reputational gain from attending an elite school for the marginal individual (ability  $\alpha^{EL}$ ). The right hand side is the loss in skill acquisition due to lower effort and travel costs. When these are equal, students with ability  $\hat{\alpha} > \alpha^{EL}$  strictly prefer being at a selective school, while students with lower ability are better off at non-selective schools with a distribution of abilities given by a normal truncated at  $\alpha^{EL}$ . Given that the reputational gain at the margin goes to zero with the entry of selective schools ( $\lim_{\alpha^{EL} \rightarrow -\infty} (\alpha^{EL} - Z(\alpha^{EL})) - g(\alpha^{EL}) = 0$ ), it follows that if there is any cost, no matter how small, from attending a selective school then a hybrid equilibrium exists.

A natural question is how such an equilibrium changes with changes in transportation costs (or equivalently an increase in population density). We can sign the effect of transportation costs upon the most selective hybrid equilibrium:

**Proposition 9.** *Suppose that transportation costs are strictly positive, then there is a stable hybrid equilibrium characterized by  $\alpha^{EL*}(tc^{EL})$ , the smallest cutoff for which such an equilibrium exists. This equilibrium has the feature that  $\alpha^{EL*}$  is increasing with transportation costs.*

This result provides a potential explanation for the fact that cities like New York and Los Angeles display low educational performance for low socioeconomic status students, despite having dense, highly competitive school markets and high expenditure. Their high density lowers transportation costs, leading to increased stratification, reduced incentives for students in non-selective schools, and reduced incentives for schools to differentiate on productivity.<sup>34</sup> This is consistent with the persistent inequality and segregation observed in these settings (Cutler and Glaeser (1997)).

## 6. EXAM SCHOOLS

Many jurisdictions respond to a demand for better schools by introducing *exam* schools. These are selective schools with test based admissions, as discussed in Section 2. In our framework their introduction can *worsen* performance. First, as with elite schools, the cream-skimming of non-selective schools by exam schools lowers the relative diversity faced by all students, reducing effort. Second, if the size of the exam sector is limited, then the most able students will strictly prefer to enroll in it, regardless of exam schools' true value added. This prediction is consistent with recent results in Abdulkadiroglu, Angrist, and Pathak (2011) and Dobbie and Fryer (2011). These papers find that students who just make it into exam schools such as Boston Latin (Boston) or Bronx Science (New York) do not obtain higher test scores than those who just miss gaining admission. Thus, the tremendous excess demand for slots at these prestigious institutions cannot be explained in terms of superior value added, but is consistent with a desire by students to attend a school with a high reputation.

## 7. DISCUSSION

When Friedman (1962) argued for the importance of reputation effects in the market for education, his goal was *normative*—he believed that competition and schools' desire for good reputations would result in

---

<sup>34</sup> As Bénabou (1993) shows, travel costs are indeed central to segregation—the benefit of a city is that it allows a large population to be close to employment opportunities, which in turn allows for segregation.

more productive schools. We show that the introduction of reputation effects plays an important *positive* role—it can help explain a set of empirical facts that at first seem unrelated to each other.

Our model has two basic ingredients. We begin with Jovanovic’s (1979) insight that individual productivity is known only imperfectly, and hence the market rationally updates its beliefs as new information emerges. This implies that if schools are selective and can observe information not directly available to the market, then employers will use school identity in setting wages. This prediction is consistent with the results of Saavedra (2009) and Arcidiacono, Bayer, and Hizmo (2010). Moreover, this effect should diminish over time, a prediction that is consistent with the findings of Dale and Krueger (2002). This ingredient explains why students prefer schools with better peers even when, as recent research suggests, such schools do not necessarily supply greater value added.

The second ingredient is the effect of reputation upon incentives. Holmstrom (1999) shows that in a competitive market a more precise estimate of individual skill may lead to lower investment in skills.<sup>35</sup> We use this idea to define a school’s *relative diversity*—a measure of ability dispersion relative to the market’s information. We show that unfettered competition and choice, by reducing relative diversity, can in some cases lead to lower student effort and skill accumulation. This prediction is consistent with the finding that as U.S. colleges have become more selective, study time has declined. It is also consistent with the finding that achievement is lower in disadvantaged neighborhoods in major urban areas (Cutler and Glaeser (1997)).

These two ingredients account for our finding that unfettered competition with reputation effects results in less than first best skill acquisition. When schools can set admissions policies reputational concerns lead to an *anti-lemons* effect: excessive entry by selective schools that cream-skin the best students from non-selective schools.<sup>36</sup> This results in lower effort, and in some cases, even lower average school productivity.

The model obviously has implications for school market design. Among the feasible systems we study, the most efficient features school choice with randomized admissions, as is currently required for most charter schools in the U.S. The model also predicts that system performance should increase with the quality of the graduation test, and with the introduction of national exams in jurisdictions that do not have them. This is consistent with the recommendations of Bishop (2006) and Woessmann (2007).

We note these policy implications abstract from some potential gains that selective systems may produce. For example, there may be social value in signals of ability if they allow for better matching of specific educational inputs to students, or the matching of individuals with particular skills to corresponding jobs. Teasing out these countervailing effects is a challenging agenda for future research.

---

<sup>35</sup> See Dewatripont, Jewitt, and Tirole (1999) for an extension to more general information structures. Coate and Loury (1993) use this idea to show that affirmative action may not necessarily enhance outcomes for disadvantaged individuals.

<sup>36</sup> Analogous phenomena have been observed in other markets. For example, Dranove et al. (2003) show that health report cards can result in cream-skimming of patients by physicians, leading to under-provision of services to the most needy individuals.

In summary, our results show that the introduction of school reputation into an otherwise perfectly competitive labor market provides a surprisingly rich set of predictions that are consistent with existing evidence. Much work remains to be done. The evidence on the impact of relative diversity needs to be explored carefully. We have not addressed the issue of tuition policy, nor the role of for profits and non-profits in the market for education. Our model assumes perfect competition and does not address issues such as class and school size. A next step would be to integrate school reputation into a model with tuition and monopolistic competition along the lines of Epple and Romano (1998).

#### REFERENCES

- ABDULKADIROGLU, A., J. ANGRIST, S. DYNARSKI, T. KANE, AND P. PATHAK (2011): "Accountability and flexibility in public schools: Evidence from Boston's charters and pilots," *Quarterly Journal of Economics*, 126, 699–748.
- ABDULKADIROGLU, A., J. ANGRIST, AND P. PATHAK (2011): "The elite illusion: Achievement effects at Boston and New York exam schools," Mimeo, National Bureau of Economic Research Working Paper No. 17264.
- AKERLOF, G. A. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84(3), 488–500.
- ALTONJI, J. G., AND C. R. PIERRET (2001): "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 116(1), 313–50.
- ANGRIST, J., E. BETTINGER, E. BLOOM, M. KREMER, AND E. KING (2002): "The effect of school vouchers on students: Evidence from Colombia," *American Economic Review*, 92(5), 1535–1558.
- ARCIDIACONO, P., P. BAYER, AND A. HIZMO (2010): "Beyond Signaling and Human Capital: Education and the Revelation of Ability," *American Economic Journal: Applied Economics*, 2, 76–104.
- ARNOTT, R., AND J. G. ROWSE (1987): "Peer Group Effects And Educational-attainment," *Journal of Public Economics*, 32(3), 287–305.
- AUMANN, R. J. (1966): "Existence of Competitive Equilibria in Markets with a Continuum of Traders," *Econometrica*, 34(1), pp. 1–17.
- BABCOCK, P. S., AND M. MARKS (2010): "The Falling Time Cost of College: Evidence from Half a Century of Time Use Data," Discussion paper, NBER.
- BARROW, L., AND C. ROUSE (2009): "School vouchers and student achievement: Recent evidence and remaining questions," *Annual Review of Economics*, January(17-42).
- BÉNABOU, R. (1993): "Workings of a City: Location, Education, and Production," *The Quarterly Journal of Economics*, 108(3), 619–652.
- BIRNBAUM, Z. W. (1950): "Effect of Linear Truncation on a Multinormal Population," *The Annals of Mathematical Statistics*, 21(2), 272–279.
- BISHOP, J. (2006): "Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn," in *Handbook of the Economics of Education, Volume 1*, ed. by E. A. Hanushek, and F. Welch, vol. 2, chap. 15, pp. 909–944. Elsevier B.V.
- BLACK, S. (1999): "Do better schools matter? Parental valuation of elementary education," *The Quarterly Journal of Economics*, 114(2), 577–599.
- BÖHLMARK, A., AND M. LINDAHL (2012): "Independent Schools and Long-Run Educational Outcomes - Evidence from Sweden's Large Scale Voucher Reform," Discussion paper, Stockholm University, Swedish Institute for Social Research Stockholm University.
- BOWEN, W. G., M. M. CHINGOS, AND M. S. MCPHERSON (2009): *Crossing the Finish Line*. Princeton University Press.
- CARD, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3, pp. 1801–1863. North-Holland, Amsterdam.

- CARRELL, S. E., B. I. SACERDOTE, AND J. E. WEST (2010): “Beware of Economists Bearing Reduced Forms? An Experiment in How Not To Improve Student Outcomes,” NBER Summer Institute 2010.
- CLARK, D. (2010): “Selective schools and academic achievement,” *B.E. Journal of Economic Analysis and Policy: Advances*.
- COATE, S., AND G. LOURY (1993): “Will Affirmative-Action Policies Eliminate Negative Stereotypes?,” *American Economic Review*, 85(5), 1220–1240.
- COLEMAN, J. S. (1960): “The Adolescent Subculture and Academic Achievement,” *American Journal of Sociology*, 65(4), pp. 337–347.
- CULLEN, J., B. JACOB, AND S. LEVITT (2005): “The effect of school choice on student outcomes: An analysis of the Chicago public schools,” *Journal of Public Economics*, 89(5-6), 729–760.
- (2006): “The effect of school choice on student outcomes: Evidence from randomized lotteries,” *Econometrica*, 74(5), 1191–1230.
- CURRIE, J. (2009): “Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development,” *Forthcoming Journal of Economic Literature*, 47(1), 87–122.
- CUTLER, D. M., AND E. L. GLAESER (1997): “Are Ghettos Good or Bad?,” *The Quarterly Journal of Economics*, 112(3), pp. 827–872.
- DALE, S. B., AND A. B. KRUEGER (2002): “Estimating the Payoff to Attending A More Selective College: An Application of Selection on Observables and Unobservables,” *Quarterly Journal of Economics*, 117(4), 1491–1527.
- DANG, H.-A., AND H. ROGERS (2008): “The growing phenomenon of private tutoring: Does it deepen human capital, widen inequalities, or waste resources?,” *The World Bank Research Observer*, 23(2), 161–200.
- DEMING, D. J., J. S. HASTINGS, T. J. KANE, AND D. O. STAIGER (2011): “School choice, School Quality and Postsecondary Attainment,” Discussion Paper 17438, NBER, Cambridge, MA.
- DEWATRIPONT, M., I. JEWITT, AND J. TIROLE (1999): “The Economics of Career Concerns, Part I: Comparing Information Structures,” *REStudies*, 66(1), 183–98.
- DOBBIE, W., AND R. FRYER (2011): “Exam high schools and academic achievement: Evidence from New York City,” Working paper 17286, National Bureau of Economic Research.
- DRANOVE, D., D. KESSLER, M. MCCLELLAN, AND M. A. SATTERTHWAITE (2003): “Is More Information Better? the Effects of “Report Cards” on Health Care Providers,” *Journal of Political Economy*, 111(3), 555–587.
- EPPLE, D., AND R. ROMANO (2008): “Educational Vouchers And Cream Skimming,” *International Economic Review*, 49(4), 1395–1435.
- EPPLE, D., AND R. E. ROMANO (1998): “Competition between Private and Public Schools, Vouchers, and Peer-Group Effects,” *American Economic Review*, 88(1), 33–62.
- FARBER, H. S., AND R. GIBBONS (1996): “Learning and Wage Dynamics,” *Quarterly Journal of Economics*, 111(4), 1007–47.
- FRIEDMAN, M. (1962): *Capitalism and Freedom*. University of Chicago Press, Chicago, With the assistance of Rose D. Friedman. 23 cm.
- GABSZEWICZ, J., AND J. THISSE (1979): “Price-competition, Quality And Income Disparities,” *Journal of Economic Theory*, 20(3), 340–359.
- GIBBONS, R., AND K. J. MURPHY (1992): “Optimal Incentive Contracts in the Presence of Career Concerns,” *Journal of Political Economy*, 100, 30–52.
- GORDON, R. D. (1941): “The Estimation of a Quotient when the Denominator is Normally Distributed,” *The Annals of Mathematical Statistics*, 12(1), pp. 115–118.
- GROGGER, J. (2011): “Speech Patterns and Racial Wage Inequality,” *Journal of Human Resources*, 1(1), 1–25.
- HASTINGS, J., AND J. WEINSTEIN (2008): “Information, School Choice, and Academic Achievement: Evidence from Two Experiments,” *Quarterly Journal of Economics*, 123(4), 1373–1414.
- HATFIELD, J. W., F. KOJIMA, AND Y. NARITA (2011): “Promoting School Competition Through School Choice: A Market Design Approach,” mimeo, Stanford University.
- HIRSCH, M. W., AND S. SMALE (1974): *Differentiable Equations, Dynamical Systems, and Linear Algebra*. Academic Press, San Diego, CA.

- HOEKSTRA, M. (2009): “The effect of attending the flagship state university on earning: A discontinuity approach,” *Review of Economics and Statistics*, 91(4), 717–724.
- HOLMSTROM, B. (1999): “Managerial incentive problems: A dynamic perspective,” *Review of Economic Studies*, 66(1), 169–182.
- HOLMSTRÖM, B., AND R. MYERSON (1983): “Efficient and Durable Decision Rules with Incomplete Information,” *Econometrica*, 51, 1799–1819.
- HOXBY, C. M. (2009): “The Changing Selectivity of American Colleges,” *Journal of Economic Perspectives*, 23(4), 95–118.
- HSIEH, C.-T., AND M. URQUIOLA (2006): “The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile’s School Voucher Program,” *Journal of Public Economics*, 90, 1477–1503.
- JACKSON, C. K. (2010): “Do students benefit from attending better schools?: Evidence from rule based student assignments in Trinidad and Tobago,” *The Economic Journal*.
- JOVANOVIC, B. (1979): “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87, 972–90.
- KATZ, L., J. KLING, AND J. LIEBMAN (2006): “Experimental Analysis of Neighborhood Effects,” *Econometrica*, 75(1), 83–119.
- KRUEGER, A., AND P. ZHU (2004): “Another look at the New York City voucher experiment,” *Behavioral Scientist*, January, 658–698.
- LANGE, F. (2007): “The Speed of Employer Learning,” *Journal of Labor Economics*, 25(1-35), 497–532.
- LEMANN, N. (1999): *The big test: The secret history of American meritocracy*. Farrar, Strauss, and Giroux, New York, NY.
- MACLEOD, W. B. (2007): “Reputations, Relationships and Contract Enforcement,” *Journal of Economics Literature*, XLV, 597–630.
- MACLEOD, W. B., AND M. URQUIOLA (2009): “Anti-Lemons: School reputation and educational quality,” Discussion Paper 15112, National Bureau for Economics Research.
- MCEWAN, P., M. URQUIOLA, AND E. VEGAS (2008): “School choice, stratification, and information on school performance: Lessons from Chile,” *Economia*, 8(2), 1–27.
- MILGROM, P. (1981): “Good News and Bad News: Representation Theorems and Applications,” *Bell Journal of Economics*, 12, 280–312.
- MINCER, J. (1974): *Schooling, Experience and Earnings*. Columbia University Press, New York, NY.
- MIZALA, A., AND M. URQUIOLA (2008): “Parental choice and school markets: The impact of information on school effectiveness,” Mimeo, Columbia University.
- NEAL, D. (2010): “The role of private schools in education markets,” in *Handbook of Research on School Choice*, ed. by M. Berends, M. G. Springer, D. Ballou, and H. J. Walber. Lawrence Erlbaum Associates/Taylor & Francis Group.
- OREOPOULOS, P. (2003): “The Long-run Consequences of Living in a Poor Neighborhood,” *Quarterly Journal of Economics*, 118(4), 1533–1575.
- OREOPOULOS, PHILIP WACHTER, T. V. H. A. (2006): “The Short and Long-term Career Effects of Graduating in a Recession: Hysteresis and Heterogeneity In the Market for College Graduates,” Discussion Paper 12159, NBER, Cambridge, MA.
- POP-ELECHES, C., AND M. URQUIOLA (2011): “The consequences of going to a better school,” Mimeo, Columbia University.
- (forthcoming): “Going to a better school: Effects and behavioral responses,” Mimeo, American Economic Review.
- ROTHSTEIN, J. (2006): “Good Principals or Good Peers: Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions,” *American Economic Review*, 96(4), 1333–1350.
- SAAVEDRA, J. E. (2009): “The Role of Resources and Incentives in Education Production,” Ph.D. thesis, Harvard University.
- SOLON, G. (1999): “Intergenerational Mobility in the Labor Market,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3, chap. 29, pp. 1761–1800. North-Holland, Amsterdam.



- TURNER, L. (2012): “The incidence of student financial aid: Evidence from the Pell grant program,” Mimeo, Columbia University.
- WOESSMANN, L. (2007): “International Evidence on School Competition, Autonomy, and Accountability: A Review,” *Peabody Journal of Education*, 82(2-3), 473–497.

## APPENDIX A. PROOFS

### Proof of Proposition 1:

Proposition 1 follows from: i) our assumptions regarding the distribution of innate ability, and ii) equations (2.1), (2.2), and (2.3) for non-selective, exam, and elite schools, respectively.

### Proof of Proposition 2:

Consider the non-selective schools first. The market estimates skill,  $\theta_i$ , using two signals,  $t_i$  and  $R_{s_i}$ . When schools are non-selective, the expected skill depends only upon the prior and hence  $E\{\theta_i|R_{s_i}\} = \hat{\alpha}_s^{NS} + \hat{e}_s + v_s$ , with precision  $\rho$ . The graduation test  $t$  is a direct measure of skill, with precision  $\rho^t$ . Hence, Bayesian learning implies column 2 of Table 2. The remaining cases are similar, and follow from the fact that for exam schools we have  $E\{\theta_i|R_{s_i}\} = \hat{\alpha}_s^{EX} + \hat{e}_s + v_s$  with precision  $\rho^\tau + \rho$ ; for elite schools we have  $E\{\theta_i|R_{s_i}\} = \hat{\alpha}_s^{EL} + \hat{e}_s + v_s$  with precision  $\rho^b + \rho^\tau + \rho$ .

### Proof of Proposition 3:

We use log linear payoffs repeatedly below. The preferences of individual  $\psi$  at school  $s$  are given by:

$$(A.1) \quad u(\psi, s) = (\gamma_1 e_{\psi s} - \gamma_0 d(e_{\psi s})) + (\gamma_1 v_s - \gamma_0 m) + \tilde{u}(\psi),$$

where the first term depends only upon effort, the second term only upon the tax rate and allocation to schools, and the third term is independent of the allocation. Given this it follows immediately that condition 1 is necessary for any efficient allocation.

The second term is independent of student characteristics, hence it is immediate that at any efficient allocation students should attend only high productivity schools, which implies condition 2. Hence, the only issue is the level of tax. To maximize net value added we need to solve for  $s \in S_H$ :

$$\max_m (-\gamma_0 m + \gamma_1 \bar{K} (m \bar{Y} Q_H)),$$

which implies condition 3.

Given that the third term in preferences is independent of the allocation, and that each student is choosing the same optimal parameters, then these conditions are also sufficient.

### Proof of Proposition 4:

The first step is to observe that in equilibrium all schools that are not oversubscribed must give the same utility to each type, which in turn implies that they must have the same relative diversity—otherwise a large mass would wish to switch to these schools. Next if they all have the same relative diversity, then students prefer the school with the highest reputation, and hence they must have the same reputation. But given that over subscribed schools randomly select students, the distribution of students who are not admitted to an oversubscribed school is the same as the distribution of those who are. Thus it must be the case that  $\hat{\alpha}_s = 0$  and  $\pi_s = \pi^{NS}$ .

Given that the off-equilibrium beliefs for schools with no students entails no selection, then if a low productivity school has students it will always be out-competed by a high productivity school. Since high productivity schools can supply the whole market, then low productivity schools have no students. This implies that the tax rate is set at  $m^H$ .

From Proposition 2, the market measures skill via the graduation test and school reputation:

$$\begin{aligned} w^{NS}(s_i, t_i) &= \pi^{NS} t_i + (1 - \pi^{NS}) R_{s_i} \\ &= \pi^{NS} (t_i - \hat{e}_s - v_s) + \hat{e}_s + v_s. \end{aligned}$$

Thus, the payoff of a student with characteristics  $\psi$  who enters school  $s$  is:

$$\begin{aligned} u_{\psi s}(c_{\psi s}^0, c_{\psi s}^1, e_{\psi s}) &= \gamma_0 \{y_\psi - m - d(e_{\psi s})\} \\ &+ \gamma_1 E \left\{ \pi^{NS}(t_{\psi s} - \hat{e}_s - v_H(m)) + \hat{e}_s + v_H(m) \mid \psi \right\}. \end{aligned}$$

From this expression one has that utility is additively separable in effort, and hence the incentive compatibility constraint implies:

$$d'(e^{PC}) = \pi^{NS} \frac{\gamma_1}{\gamma_0},$$

and hence  $e^{PC} = e^{NS}$ , as in the benchmark non-selective system (Section 2).

### Proof of Proposition 5

For individual  $i$  observe that  $(\alpha_i, \hat{\alpha}_i, t_i)$  has a multivariate normal distribution with strictly positive covariance matrix and zero means. Let  $f(\alpha_i, \hat{\alpha}_i, t_i)$  be the corresponding probability density function, and  $f(\hat{\alpha}_i, t_i)$  the density when  $\alpha_i$  is integrated out. This implies:

$$E \{ E \{ \alpha_i \mid \hat{\alpha}_i \in [\underline{\alpha}_s, \bar{\alpha}_s], t_i \} \mid \hat{\alpha}_i \} = \frac{\int_{-\infty}^{\infty} \hat{\alpha}(t_i, s) f(\hat{\alpha}_i, t_i) dt_i}{\int_{-\infty}^{\infty} f(\hat{\alpha}_i, t_i) dt_i}$$

where:

$$\begin{aligned} \hat{\alpha}(t_i, s) &= \left\{ \frac{\int_{\underline{\alpha}_s}^{\bar{\alpha}_s} \int_{-\infty}^{\infty} \alpha f(\alpha, \hat{\alpha}, t_i) h(\hat{\alpha}) d\alpha d\hat{\alpha}}{f_s(t_i)} \right\}, \\ f_s(t_i) &= \int_{\underline{\alpha}_s}^{\bar{\alpha}_s} \int_{-\infty}^{\infty} f(\alpha, \hat{\alpha}, t_i) h(\hat{\alpha}) d\alpha d\hat{\alpha}. \end{aligned}$$

is the expected ability of an individual who attended school  $s$  and has graduation test score  $t_i$ . When computing her future wage, the student has to take into account the correlation between  $\hat{\alpha}_i$  and  $t_i$ . When selecting a school the student computes the expected utility given the future expected distribution of graduation test scores  $t_i$ . If for each realization of  $t_i$  the effect of increasing selectivity is positive, then so will be the effect of selectivity upon expected future wages at the time of admission.

Thus we have:

$$\begin{aligned} \frac{\partial \hat{\alpha}(t_i, s)}{\partial \underline{\alpha}_s} &= \frac{f_s(\underline{\alpha}_s, t_i)}{f_s(t_i)} \left\{ - \frac{\int_{-\infty}^{\infty} \alpha f(\alpha, \underline{\alpha}_s, t_i) h(\underline{\alpha}_s) d\alpha}{f_s(\underline{\alpha}_s, t_i)} + \hat{\alpha}(t_i, s) \right\} \\ &= \frac{f_s(\underline{\alpha}_s, t_i)}{f_s(t_i)} \{ \hat{\alpha}(t_i, s) - E \{ \alpha_i \mid t_i, \hat{\alpha}_i = \underline{\alpha}_s \} \} > 0. \end{aligned}$$

where

$$f_s(\underline{\alpha}_s, t_i) = \int_{-\infty}^{\infty} f(\alpha, \underline{\alpha}_s, t_i) h(\underline{\alpha}_s) d\alpha.$$

is the marginal distribution given  $(\underline{\alpha}_s, t_i)$ . The final inequality follows from  $h(\underline{\alpha}_s) > 0$ , and from the fact that  $\hat{\alpha}_i$  is a good signal for  $\alpha_i$  in the sense of Milgrom (1981).

### Proof of Claims in Section 5.3

By definition we have:

$$\hat{\alpha} = E \{ \alpha \mid \hat{\alpha} \}.$$

Let  $G$  be the  $\sigma$ -algebra generated by the realizations of  $\hat{\alpha}$  and define the  $\sigma$ -algebra  $G(\bar{\alpha}) = \{ \hat{\alpha} \leq \bar{\alpha} \}$ . Since  $G(\bar{\alpha}) \subset G$  we have the following equations from the properties of conditional expectations, the fact that  $\hat{\alpha}$

has full support, and  $E\{\hat{\alpha}\} = 0$ :

$$\begin{aligned} Z(\bar{\alpha}) = E\{\alpha|\hat{\alpha} \leq \bar{\alpha}\} &= E\{\alpha|G(\bar{\alpha})\} \\ &= E\{E\{\alpha|G\}|G(\bar{\alpha})\} \\ &= E\{\hat{\alpha}|G(\bar{\alpha})\} \\ &= E\{\hat{\alpha}|\hat{\alpha} \leq \bar{\alpha}\} \\ &< \min\{\bar{\alpha}, 0\}. \end{aligned}$$

This implies that  $Z(\bar{\alpha})$  is monotonically increasing with  $\bar{\alpha}$  and

$$\lim_{\bar{\alpha} \rightarrow \infty} Z(\bar{\alpha}) = 0.$$

This expected ability can be explicitly computed using results on conditional expectations of normal random variables with truncation. From Birnbaum (1950) we have:

$$E\{X|Q \geq q\} = \mu R(q),$$

where  $X$  and  $Q$  are standard normal random variables with variances  $\sigma_X^2$  and  $\sigma_Q^2$ , respectively;  $\mu = E\{XQ\}$ ,  $R(q) = \frac{f(q)}{1-F(q)}$  is the inverse Mills ratio, and  $f$  and  $F$  are the p.d.f. and c.d.f. for the standard normal distribution. If  $X$  and  $Q$  have normal distributions, then:

$$(A.2) \quad E\{X|Q \leq q\} = E\{X\} - \frac{\text{cov}\{X, Q\}}{\sigma_Q} R\left(\frac{E\{Q\} - q}{\sigma_Q}\right).$$

Applying this to  $\alpha$  and  $\hat{\alpha}$  we get:

$$(A.3) \quad Z(\bar{\alpha}) = -\frac{\sigma_{\alpha, \hat{\alpha}}^2}{\sigma_{\hat{\alpha}}} R\left(-\frac{\bar{\alpha}}{\sigma_{\hat{\alpha}}}\right),$$

where  $\sigma_{\hat{\alpha}}^2 = \text{var}(\hat{\alpha})$  and  $\sigma_{\alpha, \hat{\alpha}}^2 = \text{cov}(\alpha, \hat{\alpha})$ . Since

$$\hat{\alpha} = E\{\alpha|\hat{\alpha}\},$$

then:

$$\alpha = \hat{\alpha} + \epsilon,$$

where  $\epsilon$  is normally distributed and uncorrelated with  $\hat{\alpha}$ ; this implies by the Gauss Markov theorem that:

$$1 = \frac{\text{cov}(\alpha, \hat{\alpha})}{\sigma_{\hat{\alpha}}^2},$$

which implies  $\sigma_{\hat{\alpha}}^2 = \text{cov}(\alpha, \hat{\alpha})$ , from which we can derive:

$$(A.4) \quad Z(\alpha^{EL}) = E\{\alpha|\hat{\alpha} \leq \alpha^{EL}\} = -\sqrt{\text{cov}(\alpha, \hat{\alpha})} R\left(-\frac{\alpha^{EL}}{\sqrt{\text{cov}(\alpha, \hat{\alpha})}}\right),$$

where  $R(z) = \frac{f(z)}{1-F(z)}$  is the inverse Mills ratio.

From Gordon (1941), section 3, page 11, we have for  $x > 0$ :

$$(A.5) \quad x + \frac{1}{x} > R(x) > x.$$

This expression and (A.3) together imply for  $\bar{\alpha} < 0$  expression 5.3.

### Proof of Proposition 7

Let  $\bar{s}$  denote a school in the non-selective sector characterized by having students with ability  $\hat{\alpha} \leq \alpha^{EL}$ . Let  $w(t, \alpha^{EL})$  be the expected wage of an individual from the non-selective sector with equilibrium cutoff  $\alpha^{EL}$  who enters the labor market with a graduation test score  $t$ . For student  $i$  at this school  $\frac{\partial t_{i\bar{s}}}{\partial e_{i\bar{s}}} = 1$  and

thus the level of effort,  $e_{\bar{s}}(\hat{\alpha}, \alpha^{EL})$ , in the non-selective sector solves:

$$= -d'(e_{\bar{s}}(\hat{\alpha}, \alpha^{EL})) = \frac{\gamma_1}{\gamma_0} E \left\{ \frac{\partial w(t, \alpha^{EL})}{\partial t} \Big| \hat{\alpha} \right\} = \frac{\gamma_1}{\gamma_0} E \left\{ \frac{\partial E\{\alpha|t, \hat{\alpha} \leq \alpha^{EL}\}}{\partial t} \Big| \hat{\alpha} \right\}.$$

We assume that in equilibrium value added and study time are known, and hence we can substitute  $t^\alpha = t - v_{\bar{s}} - \hat{e}_{\bar{s}}$  for  $t$ , and observe that  $\frac{\partial t_{i\bar{s}}}{\partial e_{i\bar{s}}} = \frac{\partial t_{i\bar{s}}^\alpha}{\partial e_{i\bar{s}}}$ .

We can use (A.2) to compute  $E\{\alpha|t^\alpha, \hat{\alpha} \leq \alpha^{EL}\}$  by letting  $X(t^\alpha) = E\{\alpha|t^\alpha, \hat{\alpha}\}$  and taking all expectations conditional upon  $t^\alpha$ , we have:

$$\begin{aligned} E\{\alpha|t^\alpha, \hat{\alpha} \leq \alpha^{EL}\} &= E\{E\{\alpha|t^\alpha, \hat{\alpha}\} | t^\alpha, \hat{\alpha} \leq \alpha^{EL}\}, \\ &= E\{X(t^\alpha) | t^\alpha, \hat{\alpha} \leq \alpha^{EL}\}. \end{aligned}$$

Using the Bayesian learning rule:

$$X(t^\alpha) = \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t} t^\alpha + \frac{\rho^b + \rho^\tau}{\rho + \rho^b + \rho^\tau + \rho^t} \hat{\alpha},$$

where  $\hat{\alpha} = \alpha + \hat{\epsilon}$ , where  $\hat{\epsilon}$  is normally distributed with mean zero and precision  $\rho^b + \rho^\tau$ . From these expressions we compute:

$$\begin{aligned} E\{X(t^\alpha)\} &= \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t} t^\alpha + \frac{\rho^b + \rho^\tau}{\rho + \rho^b + \rho^\tau + \rho^t} E\{\hat{\alpha}|t^\alpha\} \\ &= \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t} t^\alpha + \frac{\rho^b + \rho^\tau}{\rho + \rho^b + \rho^\tau + \rho^t} \frac{\rho^t}{\rho + \rho^t} t^\alpha \\ &= \frac{\rho^t}{\rho + \rho^t} t^\alpha. \end{aligned}$$

Let:

$$\begin{aligned} E\{\hat{\alpha}|t^\alpha\} &= \frac{\rho^t}{\rho + \rho^t} t^\alpha, \\ \text{var}\{\hat{\alpha}|t^\alpha\} &= \text{var}(\alpha|t^\alpha) + \text{var}(\hat{\epsilon}), \\ &= \frac{1}{\rho + \rho^t} + \frac{1}{\rho^b + \rho^\tau}, \\ &= \frac{\rho + \rho^b + \rho^\tau + \rho^t}{\rho^b + \rho^\tau (\rho + \rho^t)} \\ &\equiv \hat{\sigma}_t^2 \end{aligned}$$

Finally the covariance:

$$\text{cov}(X(t^\alpha), E\{\hat{\alpha}|t^\alpha\}) = \frac{\rho^b + \rho^\tau}{\rho + \rho^b + \rho^\tau + \rho^t} \text{var}(\hat{\alpha}|t^\alpha) = \frac{1}{\rho + \rho^t}.$$

We now use formula (A.2) to obtain:

$$E\{\alpha|t^\alpha, \hat{\alpha} \leq \alpha^{EL}\} = \frac{\rho^t}{\rho + \rho^t} t^\alpha - \frac{1}{\hat{\sigma}_t (\rho + \rho^t)} R \left( \frac{\frac{\rho^t}{\rho + \rho^t} t^\alpha - \alpha^{EL}}{\hat{\sigma}_t} \right).$$

We can now determine the marginal effect of study effort for student  $i$  in the non-selective sector upon future wages:

$$(A.6) \quad \frac{\partial w(t_i, \alpha^{HS})}{\partial e_i} = \frac{\rho^t}{\rho + \rho^t} \left( 1 - \frac{\rho^b + \rho^\tau}{(\rho + \rho^b + \rho^\tau + \rho^t)} R' \left( \frac{\frac{\rho^t}{\rho + \rho^t} t_i^\alpha - \alpha^{EL}}{\hat{\sigma}_t} \right) \right),$$

where  $t_i^\alpha$  is student's  $i$ 's signal less equilibrium effort and value added. From the inequalities (A.5) on  $R$  and the fact that  $R'(x) = R(x)(R(x) - x)$  one can show that  $R'' > 0$  it follows that  $1 > R' > 0$ , from which we

conclude:

$$\frac{\rho^t}{\rho + \rho^t} > \frac{\partial w(t_i, \alpha^{EL})}{\partial e_i} > \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t},$$

and therefore:

$$\frac{\rho^t}{\rho + \rho^t} > E \left\{ \frac{\partial w(t_i, \alpha^{EL})}{\partial e_i} \Big| \hat{\alpha} \right\} > \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t},$$

Notice as well from (A.6) that:

$$\frac{\partial^2 w(t_i, \alpha^{EL})}{\partial e_i \partial \alpha^{EL}} > 0$$

and we get that the marginal return from study time rises with  $\alpha^{EL}$ .

Finally, given the properties of the unconditional expectation and (A.5) it also the case that:

$$(A.7) \quad \lim_{\hat{\alpha} \rightarrow \infty} E \left\{ \frac{\partial w(t_i, \alpha^{EL})}{\partial e_i} \Big| \hat{\alpha} \right\} = \frac{\rho^t}{\rho + \rho^t}.$$

$$(A.8) \quad \lim_{\hat{\alpha} \rightarrow -\infty} E \left\{ \frac{\partial w(t_i, \alpha^{EL})}{\partial e_i} \Big| \hat{\alpha} \right\} = \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t}.$$

### Proof of Proposition 8

Let  $\alpha^{EL}$  denote the equilibrium cutoff satisfying (5.5). From proposition 7 we know that when the population is restricted to students with  $\hat{\alpha} \geq \alpha^{EL}$  then schools and students are making optimal choices. We have two addition conditions to check:

1. Students from the elite sector do not wish to move to the non-selective sector.

2. A high productivity school cannot attract students by choosing to be perfectly selective with  $\hat{\alpha}_s < \alpha^{EL}$ .

If a student with ability  $\hat{\alpha}_s > \alpha^{EL}$  chooses a non-selective school her payoff is given by (5.1) with  $\pi_s = \pi^{NS}(\alpha^{EL})$  and  $R_s = Z(\alpha^{EL}) + v_H(m^H) + e^{NS}(\alpha^{EL})$ . Notice that it follows that that  $\partial u / \partial \hat{\alpha}_s = \gamma_1 \pi_s < \gamma_1$ . However, the person's payoff,  $u^{EL}$ , when matched to an elite school satisfies  $\partial u^{EL} / \partial \hat{\alpha}_s = \gamma_1$ , and hence we have single crossing in types and all types  $\hat{\alpha}_s > \alpha^{EL}$  would prefer elite schools.

The same single crossing argument implies that students with  $\alpha_s < \alpha^{EL}$  would be better off at a non-selective school, and hence they would not choose elite schools if offered the possibility. The fact that only high productivity schools enter, and the tax rate is set at  $m^H$  follows from our arguments above. Thus,  $\alpha^{EL}$  defines a hybrid equilibrium.

Conversely, suppose there is a hybrid equilibrium characterized by a cutoff  $\alpha^{EL}$ . Note that

$$(1 - \pi^{NS}(\alpha^{EL})) (\alpha^{EL} - Z(\alpha^{EL})) - g(\alpha^{EL}) - tc^{EL}$$

is a continuous function of  $\alpha^{EL}$ . Hence, if

$$(1 - \pi^{NS}(\alpha^{EL})) (\alpha^{EL} - Z(\alpha^{EL})) > g(\alpha^{EL}) + tc^{EL},$$

then a student of type slightly lower than  $\alpha^{EL}$  would prefer to be in a selective school, and hence  $\alpha^{EL}$  does not define an equilibrium. Similarly, if  $(1 - \pi^{NS}(\alpha^{EL})) (\alpha^{EL} - Z(\alpha^{EL})) < g(\alpha^{EL}) + tc^{EL}$ , then a student with type slightly higher than  $\alpha^{EL}$  would prefer to be in the non-selective sector. Thus (5.5) is necessary for the existence of a hybrid equilibrium.

### Proof of Proposition 9

Let  $m(\alpha) = (\alpha - Z(\alpha)) - g(\alpha)$ . We know that  $m(\alpha)$  is continuous,  $\lim_{\alpha \rightarrow -\infty} m(\alpha) = 0$  and  $\lim_{\alpha \rightarrow \infty} m(\alpha) = \infty$ . which implies via the intermediate value theorem the existence of a unique  $\alpha^{EL}(tc) > -\infty$  solving:

$$\alpha^{EL}(tc^{EL}) = \min \{ \alpha | m(\alpha) \geq tc^{EL} \}.$$

The continuity of  $m(\cdot)$  ensures that  $m(\alpha^{EL}(tc^{EL})) = tc$ .

Let  $tc > tc'$  then one has that:

$$\{\alpha | m(\alpha) \geq tc\} \subset \{\alpha | m(\alpha) \geq tc'\},$$

which implies:

$$\begin{aligned} \min \{\alpha | m(\alpha) \geq tc\} &\geq \min \{\alpha | m(\alpha) \geq tc'\}, \\ \alpha^{EL}(tc) &\geq \alpha^{EL}(tc') \end{aligned}$$

Hence  $\alpha^{EL}(tc^{EL})$  is increasing with  $tc^{EL}$ .