

ANTI-LEMONS: SCHOOL REPUTATION, RELATIVE DIVERSITY, AND EDUCATIONAL QUALITY

W. BENTLEY MACLEOD

MIGUEL URQUIOLA

COLUMBIA UNIVERSITY AND NBER

MAY 18, 2011

ABSTRACT. The ability of firms to acquire reputations for quality is a key ingredient for the efficient provision of complex commodities in a market economy. We build a model in which students with different innate abilities acquire skills as a function of 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 use an entrance exam to select students, then competition leads to stratification by ability, reduced student effort, and in some cases lower school productivity.

JEL codes: H2, H4, I21, J31

Keywords: education, reputation, competition, labor markets.

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 thank our research assistants, Elliot Ash and Uliana Logina, for excellent work and gratefully acknowledge the support of the International Growth Center and the Russell Sage Foundation. MacLeod: bentley.macleod@columbia.edu; Urquiola: miguel.urquiola@columbia.edu.

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

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. Friedman (1962) hypothesized that this ingredient is sufficient, namely, that sellers' concern for their reputation ensures that an unfettered market is efficient.¹ In this paper, we study the market for educational services and show that this hypothesis is true only under appropriate conditions.

We build a model in which students with different innate abilities acquire skills as a function of two factors: the productivity of the schools they attend, and their own effort. We show that if schools cannot select students based upon their innate ability, then a free market tends to raise school productivity. However, if schools use an entrance exam to select students, then competition leads to stratification by ability, reduced student effort, and in some cases lower school productivity.

These 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) result 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 the *buyers* who purchase its services (as judge Scalia suggests in the first quote above). This results in a tendency for selective schools to drive non-selective ones from the market.

Analogous phenomena are observed in other markets for service goods. For example, restaurants, social clubs, and law firms are 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 an incentive to learn.

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

This matters because Holmstrom (1999) has shown that in the presence of reputation effects, this incentive depends upon the existence of uncertainty regarding ability.² When uncertainty is high, individuals work hard to show their ability to the market. Our setup captures this incentive through a concept we label *relative diversity*: the extent to which individuals can convey their ability through individual-specific measures of learning as opposed to school membership. This concept matters because free entry by schools leads—via the anti-lemons effect—to stratification by ability, and therefore to lower relative diversity and effort.

Such a link between stratification and effort is consistent with anecdotal and circumstantial evidence on student behavior. For example, students in Japan exert high effort to get into elite schools like the University of Tokyo, yet those who are successful are said to dramatically lower their effort once there. In the U.S., the selectivity of colleges has increased over the past decades (Hoxby (2009)). Babcock and Marks (2010) document that roughly during this same period, the average amount of time college students spend studying has declined significantly.

The anti-lemons effect further implies that parents prefer schools with better reputations, and value better peers *per se*. Thus, they prefer higher achieving schools even if these do not offer higher value added or positive peer effects. This can explain three empirical findings in the education literature: i) parents prefer higher achieving schools,³ ii) while selective schools almost always have higher testing performance than non-selective schools, they do not display a consistent advantage in terms of value added,⁴ and iii) the magnitude of peer effects is modest and unstable.⁵

Finally, the anti-lemons effect has several policy implications. For example, it implies that 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; to illustrate, the second quote above would be consistent with the introduction of the SAT test leading to greater stratification in U.S. higher education (see Lemann (1999)). The bottom line is that policy may be able to affect incentives and educational performance, albeit in a more complicated way than Friedman (1962) had in mind.

Our agenda is as follows. Section 2 sets out a framework in which individuals choose among schools, which in turn make two choices: how much value added to supply, and whether to base their admissions on students' performance in an admissions assessment. Only students and schools (and not employers) directly observe this measure of innate ability. Upon graduation, students enter a perfectly competitive labor market

² This model is widely used in labor economics. See for example Gibbons and Murphy (1992), Farber and Gibbons (1996), and Altonji and Pierret (2001).

³ See Black (1999) and Hastings and Weinstein (2008).

⁴ See Neal (2009), Cullen, Jacob, and Levitt (2005, 2006), and Clark (2010)

⁵ See Oreopoulos (2003), Katz, Kling, and Liebman (2006), and Carrell, Sacerdote, and West (2010).

that pays them its best estimate of their skill. This estimate is based on two signals: individuals' performance on a graduation assessment, and the reputation of the school they attended, where a school's reputation is simply the expected skill of its graduates.⁶ We assume that no individual can single-handedly change her schools' reputation (e.g., a school's reputation will reflect the skill of many cohorts of graduates).

Section 3 introduces student and school preferences, and characterizes efficient allocations. Section 4 derives equilibrium market allocations. It begins with a benchmark public system in which students are randomly assigned to schools, and shows that its performance is suboptimal relative to the efficient allocation. It then analyzes whether and how performance improves as one adds elements of choice and competition. Section 5 discusses some 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 labor market wages. We do not derive the full equilibrium here (see Section 4); rather we set out initial ingredients that illustrate several key results and testable implications.

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

- (1) In a 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. One such signal is *family background*, observed by all agents. Another is an *admissions assessment* observed by individuals and schools prior to enrollment. This measure is soft information that is not verifiable by employers.
- (2) Schools select admissions policies that, to different extents, exploit the two signals of innate ability: family background and the admissions assessment. Schools also set their educational value added. School admissions policies and levels of value added are public information. Students choose among the schools that have offered them admission. (We expand upon this step significantly in Section 4).
- (3) Students make consumption choices and decide how much effort to allocate to study. 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) an individual-specific measure of skill we term a *graduation assessment*, and ii) the identity of the school attended by each student. The labor market sets wage equal to expected skill given these signals. Individuals then choose their second period consumption.

⁶ More broadly, in a competitive labor market employers have incentives to learn about ability and use all available signals of skill, including socioeconomic background. For instance, Farber and Gibbons (1996) provide evidence that employers learn about ability over time, and Altonji and Pierret (2001) and Grogger (2009) suggest background influences wages.

The remainder of this section outlines the information structure of the model. Sections 3 and 4 provide the details regarding preferences and market entry rules.

2.2. Family background and innate ability. There is a large population of students indexed by i ; each student has exogenous log income y_i (e.g., parental income), denoted in level terms by Y_i .⁷ Log income is distributed normally:

$$y_i \sim N\left(\bar{y}, \frac{1}{\rho^y}\right)$$

where \bar{y} , is the population mean and $\rho^y = \frac{1}{\text{var}(y_i)}$ is the *precision*. Throughout we will use precision rather than variance because this simplifies the Bayesian learning expressions used below. We define student i 's family background as the deviation of her log income from the mean:

$$b_i = y_i - \bar{y}.$$

A person's background is always observable.

Student innate log ability, denoted by α_i , is potentially correlated with family background:

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

where ϵ_i^b is a normally distributed random variable with expected value of zero and precision ρ^b : $\epsilon_i^b \sim N(0, \frac{1}{\rho^b})$. When $\zeta = 0$, innate ability is uncorrelated with family background. Recent research suggests that ζ is strictly positive.⁸ Note that if a person is drawn at random from the full population, her expected innate ability is zero ($E\{\alpha_i\} = 0$) with variance:

$$\sigma^2 \equiv \frac{1}{\rho} = \zeta^2 \text{var}(b_i) + \text{var}(\epsilon_i^b) = \frac{\zeta^2}{\rho^y} + \frac{1}{\rho^b}$$

where ρ is the precision of this unconditional estimate.

2.3. The admissions assessment. We assume students undergo two assessments: one before admission to school and one as they enter the labor market. Let τ_i be the measure of innate (log) ability provided by the admissions assessment:

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

where $\epsilon_i^\tau \sim N(0, \frac{1}{\rho^\tau})$ and ρ^τ is this measure's precision. This assessment can be interpreted as an admission test, but it could also reflect information gathered through student or parental interviews. In part because it

⁷ In general we will use upper case letters to denote levels, and lower case for logs.

⁸ See Solon (1992), Cunha, Heckman, and Schennach (2010), and Almond and Currie (2011).

can include such qualitative information not easily verified by third parties, we assume it is *soft information* that is not observable to the labor market.⁹ Thus, school admission policies may exploit two signals: family background, b , which is observed by all actors, and the admissions assessment, τ , observed only by students and schools—this is the only informational imperfection in our setup.

Let student i 's characteristics be summarized by $\psi_i = \{b_i, \tau_i\}$. From the assumptions above, this has a bi-variate normal distribution with c.d.f. given by $f(\psi_i)$. The expected innate log ability of an individual given these signals is $\hat{\alpha}_i = E\{\alpha_i|\psi_i\}$. The innate ability of the students at a given school depends upon the school's admission's policy given these two signals, and upon student choice. The expected innate log ability of students in school s is given by: $\hat{\alpha}_s = E\{\alpha_i|s_i = s\}$.

2.4. Admissions policies. We consider three types of schools characterized by their admissions policies: public schools, selective public schools, and selective private schools. In this section, we consider the labor market implications when schools in a given market all have the same type. We defer to Section 4 the equilibrium analysis in which schools of different types co-exist.

By a *public school* we mean a school that accepts all students, regardless of their signal (b or τ). For now we will assume students are randomly allocated to these schools (the allocation systems that achieve this are discussed in detail in Section 4); hence for these schools expected ability is:

$$(1) \quad \hat{\alpha}_s^{NS} = E\{\alpha_i|s_i = s\} = E\{\alpha_i\} = 0$$

where the superscript NS emphasizes their non-selective nature. Although this is an extreme form of non-selectivity, it is in the spirit of public schools' admissions policies in many countries. It is also consistent with the fact that many charter schools are required to implement lotteries if oversubscribed.

Second, we consider schools that use only the admissions assessment, τ , to select students. We label these *selective public schools* to capture the fact that many educational systems feature public schools that admit students based only on testing merit—background plays no role in their selection mechanisms. We suppose the market is sufficiently thick for these schools to be completely stratified, i.e., all the students at school s have the same admissions score τ_s . This dramatically simplifies the analysis because it allows us to use a linear updating rule. By Bayes' rule that the expected innate ability of an individual from a selective public

⁹ Farber and Gibbons (1996) provide a concrete example. They show that library card ownership is correlated with ability, but is not observed when workers enter the labor market.

school s with admissions standard τ_s is:¹⁰

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

where the superscript PS emphasizes the partially selective nature of these schools (again, they exploit only one of the two signals of innate ability). In (2), $\frac{\rho^\tau}{\rho^\tau + \rho}$ is the weight assigned to τ_s . Intuitively, the more precise the admissions assessment, the greater weight is put on it as opposed to the unconditional mean of ability, $E\{\alpha_i\} = 0$. The precision of $\hat{\alpha}_s^{PS}$ is $\rho(\alpha|\tau) = \rho^\tau + \rho$, the sum of the precision of the admissions assessment and the prior precision of ability.

Finally, we consider *private schools*. The term *private* merely signals the idea that in many jurisdictions private schools can use any criteria they wish to select students, including family background. For example, they may use not just exams but also student and parental interviews. Such ingredients 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 a private school s the expected innate log ability, $\hat{\alpha}_s^{FS}$, is:

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

where the superscript FS emphasizes the fully selective nature of private schools. The weight assigned to each signal again depends upon its precision; again, the precision of $\hat{\alpha}_s^{FS}$ is the sum of that of the two signals and the prior precision of ability: $\rho(\alpha|b, \tau) = \rho^b + \rho^\tau + \rho$.

The next proposition summarizes the impact that each level of selectivity has upon the expected ability of a student at school s , and the precision of this estimate:

Proposition 1. *Suppose a student with characteristics ψ_i (background b_i and admissions assessment τ_i) attends a school s in a system featuring only one type of school. Her expected innate log ability conditional*

¹⁰ See DeGroot (1972). Harris and Holmström (1982) were the first to use this model to explore wage dynamics.

upon the type of school she enrolls in satisfies:

<i>School type:</i>	<i>Public</i>	<i>Selective Public</i>	<i>Private</i>
<i>Expected innate log ability:</i>	$\hat{\alpha}_s^{NS} = 0$	$\hat{\alpha}_s^{PS} = \frac{\rho^\tau}{\rho^\tau + \rho} \tau_s$	$\hat{\alpha}_s^{FS} = \frac{\rho^b}{\rho^b + \rho^\tau + \rho} \zeta b_i + \frac{\rho^\tau}{\rho^b + \rho^\tau + \rho} \tau_s$
<i>Precision of estimate:</i>	ρ	$\rho^\tau + \rho$	$\rho^b + \rho^\tau + \rho$
<i>Degree of school selectivity:</i>	<i>Non-Selective</i>	<i>Partially Selective</i>	<i>Fully Selective</i>

TABLE 1. Expected innate log ability by type of school system attended

The appendix includes the proof of this and all subsequent propositions.

2.5. Skill. In addition to implementing admissions policies, schools also produce educational value added. Specifically, upon graduation, the skill of a student i who attended school s is given by Θ_{s_i} . Skill depends on three factors: i) the student's innate ability, ii) her effort, and iii) her school's value added, with a multiplicative form:

$$\Theta_{s_i} = A_i E_i V_{s_i}$$

that in log terms becomes:

$$(4) \quad \log(\Theta_{s_i}) = \theta_{s_i} = \alpha_i + e_i + v_{s_i}$$

where $\alpha_i = \log(A)$ is innate log ability as defined above, $e_i = \log(E_i)$ is log student effort, and $v_{s_i} = \log(V_{s_i})$ is the log value added that student i receives at school s .¹¹ We assume a school provides the same value added, V_s , to each of its students, and that its level is observed by all actors. Finally, let e_s stand for the mean log effort level at school s , which as we show below, is uncorrelated with school value added.

2.6. Signals of skill: school reputation and the graduation assessment. The market observes the identity of the school attended by each individual and correctly anticipates school value added and school average effort. From Proposition 1, we have that school identity also provides a signal of log innate ability.

¹¹ This specification is broadly 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 (e.g., Mincer (1974) and Card (1999)). It further allows for complementarities between ability and value added.

Let a school's *reputation* be the publicly observed expected log skill of its graduates:¹²

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

We suppose throughout that an individual's effort, e_i , has no effect upon her school's reputation:

$$(5) \quad \frac{\partial R_{s_i}}{\partial e_i} = 0.$$

The intuition is that a school's reputation typically depends on the characteristics of multiple cohorts of graduates, and therefore no single individual can easily affect it. This accounts for why competition does not necessarily improve system performance in all cases.

Aside from school reputation, the market observes an individual-specific measure of skill. We label this measure a *graduation assessment* 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. Note that two aspects distinguish this measure from the admissions assessment, τ , discussed above: 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 or college entry exams in countries such as Germany, Romania, South Korea, and Turkey. 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."

Finally, since a student's graduation assessment is an individual measure of performance, it responds immediately to individual effort:

$$(6) \quad \frac{\partial t_i}{\partial e_i} = 1.$$

In summary, after admission to a school, effort does not respond to rewards associated with school reputation, but does respond to rewards linked to measures of personal performance.

2.7. The labor market. Upon leaving school, an individual with characteristics ψ_i who attended school s and obtained a graduation assessment score t_i earns a wage $W(\psi_i, s_i, t_i)$. Following Jovanovic (1979) and

¹² Strictly speaking, reputation should be written as a function. We make various assumptions regarding the availability of information, and so the functional dependence will vary with the context. The variance of this estimate is also part of school reputation, though for purposes of exposition this is made explicit only when necessary.

Harris and Holmström (1982), we assume a perfectly competitive labor market that sets an individual's wage equal to the best estimate of her skill given the signals it observes—her graduation assessment and her school's reputation:

$$W_i(\psi_i, s_i, t_i) = E \{ \Theta_i | t_i, R_{s_i} \}$$

or in log terms, as wage equations are typically specified:¹³

$$(7) \quad w_i(\psi_i, s_i, t_i) = E \{ \theta_i | t_i, R_{s_i} \}.$$

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

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's graduation assessment score is t_i and the school she attended has reputation R_{s_i} , then her log wage satisfies:*

<i>School type:</i>	<i>Public</i>	<i>Selective public</i>	<i>Private</i>
<i>Log wage:</i>	$w^{NS}(\psi_i, s_i, t_i) =$ $\pi^{NS}t_i + (1 - \pi^{NS})R_{s_i}$	$w^{PS}(\psi_i, s_i, t_i) =$ $\pi^{PS}t_i + (1 - \pi^{PS})R_{s_i}$	$w^{FS}(\psi_i, s_i, t_i) =$ $\pi^{FS}t_i + (1 - \pi^{FS})R_{s_i}$
<i>Relative diversity:</i>	$\pi^{NS} = \frac{\rho^t}{\rho^t + \rho}$	$\pi^{PS} = \frac{\rho^t}{\rho^\tau + \rho^t + \rho}$	$\pi^{FS} = \frac{\rho^t}{\rho^b + \rho^\tau + \rho^t + \rho}$

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

In Table 2 we introduce the notion of *relative diversity*. This is the precision of the graduation assessment divided by the precision of the market's estimate of student skill. In each case, relative diversity is also the weight the market places upon the graduation assessment (as opposed to school reputation) in setting wages. Going from left to right, Table 2 illustrates that (keeping the precision of the graduation assessment fixed) relative diversity is lower in systems in which schools implement more selective admissions policies.

¹³ Strictly, expression (7) should include a variance component:

$$w_i(R_{s_i}, t_i) = E \{ \theta_i | R_{s_i}, t_i \} + var \{ \theta_i | R_{s_i}, t_i \} / 2$$

Since in our setting neither student effort nor school value added affect the variance of skill, we simplify matters by viewing the variance as a second order fixed term that can be ignored. In terms of the results this has some impact upon the wage levels, but does not affect the qualitative properties of the equilibria.

Since students can affect their testing performance but not their schools' reputation (recall $\frac{\partial t_i}{\partial e_i} = 1$ and $\frac{\partial R_{s_i}}{\partial e_i} = 0$), effort increases with relative diversity. When relative diversity is higher, the extent to which an individual can affect the market's impression through the one signal she can influence—the graduation assessment—is greater, and hence so is the incentive to study. Finally, note that we use the term *relative* diversity because the impact of school selectivity depends upon both the precision of the graduation assessment and school selectivity; an increase in the precision of the graduation assessment leads to greater relative diversity and increased effort in all cases.

2.8. Discussion. The wage equations in Proposition 2 illustrate several implications of the Bayesian learning model. First, since reputation directly impacts wages, 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 Education literature:

- (1) There is clear evidence that parents prefer higher achieving 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 achieving schools produce higher *value added* is mixed. For example, while private schools frequently have higher absolute achievement than public schools, recent reviews (e.g., Neal (2008) and Barrow and Rouse (2009)) conclude this advantage does not systematically extend to value added. This is consistent with a broader literature that studies the testing impact of attending a higher achieving school or class. Here again several papers find little or no effect (e.g. Cullen, Jacob, and Levitt (2005, 2006), Clark (2010), and Duflo, Dupas, and Kremer (2008)) and some find positive effects (e.g. Pop-Eleches and Urquiola (2011) and Jackson (2010)), but no uniform pattern emerges.
- (3) The evidence on the significance and magnitude of peer effects is mixed. In this area, 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.¹⁴

Proposition 2 also implies that if effort matters for learning, then selectivity—by lowering relative diversity—will decrease skill accumulation. As stated, this is consistent with anecdotal and circumstantial evidence on student behavior. For example, students in Japan exert high effort to gain admission to elite schools like the University of Tokyo, yet those who are successful are said to dramatically lower their effort once there. In the U.S., Hoxby (2009) shows that the selectivity of colleges has increased over the past decades, and Babcock and Marks (2010) document that during this same period, the average amount of time college students spent studying declined from 40 to 27 hours per week.

This also implies that if school choice reforms are associated with stratification, they may have ambiguous effects on skill, as they may lower effort even as they raise school productivity (we model the latter effect below). Consistent with this, the most extensive implementation of a voucher program, that which took place in Chile, had little impact on average achievement even while it resulted in extensive private school entry (Hsieh and Urquiola (2006)) and one of the most stratified school systems in the world (Valenzuela et al. (2010)).

Proposition 2 also implies that introducing precise individual-specific measures of skill raises learning, as emphasized by Bishop (2006). Consistent with this, Woessmann (2007) finds 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, Romania and South Korea display extensive private tutoring industries that parents use to supplement their children’s learning at school (Dang and Rogers (2008)).

2.8.1. *Labor market implications.* Proposition 2 implies that *conditional* on years of schooling, school identity can affect wages. For instance, consider a set of individuals who attend (fully selective) private schools. Although all of them are graduates and hence have the same amount of schooling, their wages are given by

$$(8) \quad w^{FS}(\psi_i, s_i, t_i) = \pi^{FS} R_{s_i} + (1 - \pi^{FS}) t_i$$

that is, their wages vary with their schools’ reputation and their graduation assessment. This implies that the identity of the school an individual attended might be a relevant proxy for ability in a wage regression (along with other standard controls like experience, etc.).¹⁵

¹⁴ 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) and Mizala and Urquiola (2008)).

¹⁵ 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.

While few datasets 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 some other means. Expression (8) implies that an educational system—to the extent it contains selective institutions and graduation assessments—may perform this function.

A corollary is that how effectively a school system reveals ability will vary with the educational level at hand. 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; additionally, these individuals tend to come from non-selective public schools. In contrast, college graduates come from a sector in which selectivity is much more common, as are individual-specific measures (e.g., 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.

Two further sets 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 assessment. First, 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 this result is for starting wages for college graduates in Colombia; in the second for wages 7-12 years after graduation in the U.S. Second, as discussed above there is evidence of greater selectivity being correlated with reduced effort at the college level (Hoxby (2009) and Babcock and Marks (2010)).

While our model features only two periods, in practice the labor market gains information about individual ability over time, as workers gather more experience. Thus, the correlation between unobserved signals 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 also that the market’s gradually gaining more information is equivalent to the precision of the recommendation signal, t , increasing over time. If this happens, the effect of school

reputation on wages should diminish (Table 2). This is consistent with Dale and Krueger’s (2002) finding that school selectivity has little impact upon wages 20 years into workers’ careers.¹⁶

Indeed, the results in Dale and Krueger (2002) are often viewed as showing that school identity does not matter. Rather, they (along with the results on wage effects early in individuals’ careers by Saavedra (2009) and Hoekstra (2009)) are consistent with the Bayesian learning model we have used. Note also that the overall set of results is *not* consistent with a peer effects model (e.g., Epple and Romano (2008)). Such a model suggests that more selective schools must have higher value added via a peer effect. This should lead to a *permanent* increase in performance from attending a selective school.

To summarize, a disparate set of studies from education and labor economics are consistent with an application of a competitive model of wage formation with Bayesian learning.

3. PREFERENCES AND EFFICIENT ALLOCATIONS

This section defines school costs and student preferences. It also derives the efficient allocations that serve as a benchmark for comparison with scenarios in which private returns determine behavior. Throughout, we assume educational systems consist of a large number of schools, each of which contains a large number of students. We thus abstract from small numbers issues (such as class size). This yields gains in tractability and allows us to explore the implications of perfect competition.

It also differentiates our model from earlier work on school choice, such as Arnott and Rowse (1987) and Epple and Romano (1998), that focuses on monopolistic competition and markets with a relatively small number of schools. Specifically, in these models there is an efficient scale for schools, and hence each school has some market power. Conceptually, such a setup supposes that the number of schools is fixed in the short run, and that students select into schools (see Nechyba (2006) for a discussion of this class of models).¹⁷

In contrast, our concern is with understanding the implications of perfect competition in the presence of reputation effects. We therefore abstract from setup costs, and follow Aumann’s (1966) model of perfect competition in supposing that there is a continuum of schools and students. Schools can perfectly tailor their characteristics to the traits of students, and neither side has market power. Any pooling of individuals by their characteristics is endogenous, and arises only because the labor market uses school reputation to set wages.

¹⁶ 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 affect of college identity.

¹⁷ For example, the empirical section of Epple and Romano (1998) considers 4-10 schools. In contrast, many urban school systems have hundreds of schools, and as Hoxby (2009) emphasizes, many colleges face a national market.

3.1. School funding and productivity. In general we will assume schools are financed via a fixed voucher. Specifically, each school receives a per-student payment:

$$\bar{P} = T\bar{Y}$$

where \bar{Y} is mean income and T a constant tax rate.¹⁸ Thus, individual i pays a tax TY_i . We suppose that schools operate on a not for profit basis,¹⁹ such that school value added is given by the solution to:

$$\bar{P} = \frac{K(V_q)}{q}$$

where costs, K , are a function of value added, V .²⁰ School productivity is given by q_s , and takes on values of q_H and q_L for schools of high and low productivity, respectively, where $q_H = 1$ and $q_H > q_L > 0$. Thus, given a voucher $\bar{P} \geq 0$, the value added supplied by a school is either $V_H(\bar{P})$ or $V_L(\bar{P})$:

$$\bar{P} = K(V_H(\bar{P})) = \frac{K(V_L(\bar{P}))}{q_L}$$

i.e., higher productivity schools supply more value added. Recall also that $v_s(\bar{P}) = \log(V_s(\bar{P}))$ is log value added, and abuse notation to let $K(v) = K(V) = K(\exp(v))$ be the cost in terms of log value added.

Finally, we assume that school productivity is publicly observed, and that schools are not differentiated along other dimensions, like distance. Aside from its student composition, the only relevant information regarding a school is whether it is of high or low productivity.

3.2. Individual preferences. Individuals care about consumption and effort, with preferences that take a 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. Thus, individuals are risk averse, with decreasing marginal utility from income ($\gamma_0 + \gamma_1 < 1$). $D(E)$ is the dis-utility from academic effort that takes place in period 0. We assume the Bernoulli utility is given by log utility:

$$\begin{aligned} u(C^0, C^1, E) &= \log(U(C^0, C^1, E)) \\ &= \gamma_0(\log(C^0) - d(E)) + \gamma_1 \log(C^1) \end{aligned}$$

¹⁸ Making the tax level endogenous does not affect our key conclusions.

¹⁹ More generally, we could derive zero profits from free entry conditions. This is done in MacLeod and Urquiola (2009) where we show that if high productivity firms are in short supply, then they can earn profits in equilibrium.

²⁰ The cost function satisfies $K''(0) > K'(0) = K'(0) = 0$ and, for $V > 0$, $K'(V)$, $K''(V) > 0$.

where $d(E) = \log(D(E))$, $d(0) = 0$, $d' > 0$, and $d'' > 0$. Finally, for simplicity we abuse notational conventions, and define $D(e) = D(\log(e))$, where the right hand side is $D(E)$, and we let $d(e) = \log(D(\log(e)))$.²¹

We abstract from savings and borrowing, and hence individuals face period-specific budget constraints:

$$\begin{aligned} C_i^0 &\leq (1 - T)Y_i \\ C_i^1 &\leq W_i. \end{aligned}$$

3.3. The ex post efficient allocation. Given this setup, what would be a socially optimal design for a school system? What is the best that could be achieved assuming a social planner can set individual consumption and effort? To address this question, we follow Holmström and Myerson (1983) and derive the *ex post* efficient allocation. This allocation is derived at the time of school admission. At this stage students' innate ability and family background have been realized, and it is not possible to redistribute income without making some students worse off. Hence, this allocation addresses what is the best that could be achieved given that each person self-finances her education.²²

One goal of school choice is to provide incentives for the entry of high productivity schools. Thus, the efficient allocation is computed under the hypothesis that schools are of high productivity.²³ Additionally, recall that at the time of admission to school, individual characteristics consist of background and admissions assessment score: $\psi_i = \{b_i, \tau_i\}$, where these have a bi-variate normal distribution $F(\psi)$. An individual's first period consumption, effort, and school value added can be conditioned upon her characteristics, and are therefore given by $C^0(\psi_i)$, $E(\psi_i)$, and $V(\psi_i)$. Second period consumption occurs after the individual realizes her wage, and hence it may also be affected by her performance on the graduation assessment: $C^1(\psi_i, t_i)$.

Since we derive the optimal level of value added as a function of individual characteristics, we suppose that the voucher payments are also type dependent. If student i is given voucher P_i then this completely determines value added via $P_i = K(V(\psi_i))$, where $V(\psi_i)$ is the efficient level of value added for a student with characteristics ψ_i .

Given these preliminaries, the ex-post efficient allocation solves:

$$(9) \quad \max_{C_0, C_1, E, V} E\{u(C^0(\psi_i), C^1(\psi_i, t_i), E(\psi_i)) \mid \psi_i\} \quad \text{such that :}$$

$$(10) \quad C^0(\psi_i) \leq Y_i - K(V(\psi_i)) \quad \text{and}$$

²¹ Thus, lower case is consistently the log of the parameter. When we write $d'(e)$ we mean $d(\log(D(\log(e))))/de$ and so on.

²² In work available upon request, we have also derived the *ex ante* efficient allocation, which is computed *before* individuals know their type, and under the assumption that transfers can be carried out at no cost. We return to the results briefly below.

²³ It is straightforward to extend the analysis to the case in which the number of high productivity schools is constrained.

$$(11) \quad C^1(\psi_i, t_i) \leq W(\psi_i, t_i) = A_i(\psi_i, t_i) E(\psi_i) V(\psi_i)$$

The first constraint requires period 0 consumption and value added be paid out of family income. The second constraint reflects that in period 1 workers simply consume their wage, which is equal to expected skill, where $A_i(\psi_i, t_i) = E\{\exp(\alpha_i) | \psi_i, t_i\}$ is the worker's expected innate ability given her type and graduation assessment. At the optimum both constraints bind; thus this problem can be solved by substituting for consumption using the budget constraints. This allows a straightforward statement of the solution:

Proposition 3. *The ex post efficient allocation is characterized by:*

- (1) The marginal cost of effort is set equal to the rate of time preference and hence is independent of individual background and ability:

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

- (2) The efficient level of value added increases with income and ensures that the marginal cost of value added equals the rate of time preference times period 0 consumption:

$$K'(v(\psi_i)) = \frac{\gamma_1}{\gamma_0} C^0(\psi_i) = \frac{\gamma_1}{\gamma_0} ((Y_i - K(v(\psi_i)))).$$

Proposition 3 illustrates the basic trade-off between current consumption and investment in education for higher future consumption.

Finally, note that if there are low productivity schools, then the level of value added is lower. In this case, the efficient solution is to allocate students to the high productivity schools until these are fully utilized, with the remaining students allocated to the low productivity schools.²⁴ We now turn to the properties of the market equilibrium for schools as a function of the school choice rules.

4. COMPETITION: THE ANTI-LEMONS EFFECT

In this section we compare the properties of market equilibria for four school systems. First, a benchmark public system in which students are randomly assigned to schools. Second, a public system with choice—schools are non-selective and use a lottery for admissions when there is oversubscription. Third, a public school system with choice that combines selective and non-selective schools. Finally, a system in which (selective) private schools can compete with (non-selective) public schools.

²⁴ As stated, the objective of the *ex-ante* efficient allocation is to maximize the utility of individuals *before* they learn their backgrounds, and under the assumption that income redistribution is possible. The key difference in the setup is therefore that each period features a single aggregate budget constraint, as opposed to one per person. The central result is that the ex-ante efficient allocation entails equality of consumption, but not equality of effort or access to value added. More able individuals enjoy lower utility, but receive more education and work harder. This recalls Marx's slogan, "From each according to his ability, to each according to his need," and is consistent with the existence of selective public schools requiring greater effort.

The first subsection discusses the choice mechanism and how it generates an allocation of students to schools. We then discuss how the allocation affects student preferences over schools. The key ingredient here is that an individual's preference depends not only upon school value added, but also upon the allocation of the other students to schools. Finally, we compare system performance across the different cases.

4.1. The choice mechanism. Let us now fill in the details for step (2) of the timeline from Section 2.1, specifying how schools enter the market and how students choose schools.

- (1) The government (e.g. a school district) specifies the set of schools S that are free to enter, as well as the set of admissions policies that they are allowed to use. Although school productivity is observable, we assume the government cannot base the specification of S upon productivity (e.g., the public sector might have difficulty eliminating low productivity operators). Given that schools earn zero profits, they enter if and only if they expect to have some demand for their services.
- (2) Students are assigned a random ranking that is used to implement a lottery system when schools are oversubscribed. Students apply (at no cost) to any school at which they meet the admissions criteria, and provide a ranking of their choices.²⁵ In some cases, schools may require students to meet a minimum ability requirements.
- (3) Schools and students are matched, with students allocated to their most preferred school among those that accepted them.

Formally, an *allocation* is denoted by $\Gamma = \{e(\psi, s), f(\psi|s)\}_{\psi \in \Psi, s \in S}$, where $e(\psi, s)$ is the log effort chosen by students with characteristics ψ at school s . The distribution of student characteristics at school s is given by $f(\psi|s)$ and satisfies:

$$(12) \quad f(\psi) = \int_{s \in S} f(\psi|s) ds \quad \text{and} \quad \int_{\psi \in \Psi} f(\psi|s) d\psi \leq 1$$

The first condition captures the requirement that the allocation of students to schools be consistent with the aggregate distribution of types. The second condition reflects that school capacity is bounded at 1.²⁶

²⁵ These mechanisms are essentially serial dictatorships (or random serial dictatorships as defined in Pathak (2011)) and as discussed below will be strategy-proof. Analogous school assignment mechanisms are observed in reality, e.g., Pop-Eleches and Urquiola (2011). We do not formally model the discrete case and take the limit since this would add unnecessary complexity.

²⁶ A simple example would be to have $S = [0, 2]$ and all students equally distributed at each school. In that case $f(\psi) = f(\psi|s)/2$, and each school is used to half capacity. Another example would be to have students allocated based upon income and ability. In that case we have $(s^1, s^2) \in S = [0, 2]^2$. Students with ability $\tau \leq 0$ are uniformly distributed over schools with $s^1 \in [0, 1]$, and similarly for students with $\tau > 0$ over $(1, 2]$. While students with background $b \leq 0$ are uniformly distributed over $s^2 \in [0, 1]$ and similarly for students with $b > 0$. In that case for $s \in [0, 1]$ we have $f(\psi|s) = f(\psi)/4$ if $b \leq 0$ and $\tau \leq 0$ and 0 otherwise. The other three cases will be similar. This illustrates how the model can deal with many complex allocations.

4.2. **Student preferences given an allocation.** Students ultimately care about effort and consumption.²⁷

Thus, characterizing their payoffs in equilibrium requires analyzing how an allocation determines: i) the incentives for effort, and ii) the future wages students can expect at any given school. To do this, consider a student $\psi_i = \{b_i, \tau_i\}$ who attends school s under allocation Γ . Let $w(\psi_i, s, \Gamma)$ be the expected log wage should she select school s given this allocation.

An allocation defines the ability distribution at each school. Section 2 showed how a school's relative diversity, $\pi_s(\Gamma)$, affects students' marginal incentives in a log normal framework. More generally, we can define relative diversity as the marginal effect of effort. Specifically, the marginal impact of effort upon wages should the student deviate from the allocation effort $e(\psi_i, s)$ at a given school is defined by:²⁸

$$(13) \quad \pi_s(\Gamma) = \frac{\partial w(\psi_i, s, \Gamma)}{\partial e_i} = E \left\{ \frac{\partial E \{ \alpha_i | t_i, s, \Gamma \}}{\partial t_i} | \psi_i, s, \Gamma \right\}.$$

This motivates the following definition:

Definition 1. An allocation Γ is *incentive compatible* if:

$$(14) \quad d'(e(\psi_i, s)) = \pi_s(\Gamma) \frac{\gamma_1}{\gamma_0}.$$

This defines the effort at each school. Its level is independent of individual ability and income. Hence, given a distribution $f(\psi|s)$, we denote an allocation by $\Gamma = \{e_s, f(\psi|s)\}_{\psi \in \Psi, s \in S}$, assuming (14) is satisfied.

To evaluate the expected wages at a school given an allocation, a student also needs to know how enrolling there will affect her two signals of skill: school reputation and graduation assessment. Students and schools have correct expectations regarding the equilibrium allocation, and hence the allocation determines school reputation—the sum of expected student ability at the school, effort, and school value added:

$$R_s(\Gamma) = \hat{\alpha}_s + e_s + v_s.$$

Similarly, given an allocation, a student's expected performance on the graduation assessment is:

$$E \{ t_i | \psi_i, \Gamma \} = \hat{\alpha}_i + e_s + v_s.$$

²⁷ Recall $u(C^0, C^1, E) = \gamma_0(\log(C^0) - d(e)) + \gamma_1 \log(C^1)$.

²⁸ Proposition 2 works out this parameter for the cases considered below. In these, the effect of t_i is linear:

$$\frac{\partial^2 E \{ \alpha_i | s, t_i, \Gamma \}}{\partial t_i^2} = 0,$$

such that the marginal effect of effort is independent of student characteristics. Note also that relative diversity is well defined as long as there are some students attending the school.

Given these expressions, the expected utility of student ψ_i upon admission to school s under allocation Γ is:

$$(15) \quad u(\psi_i, s, \Gamma) = \gamma_0 \{ \log((1 - T) Y) - d(e_s) \} + \gamma_1 \{ \pi_s(\Gamma) (\hat{\alpha}_i + e_s + v_s) + (1 - \pi_s(\Gamma)) R_s \}.$$

Expression (15) implies:

- (1) All else equal students prefer schools with higher value added, since value added raises payoffs both through school reputation and performance on the graduation exam.
- (2) Holding relative diversity fixed, students prefer schools with better reputations.
- (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:

$$\frac{\partial u}{\partial \pi_s} = \hat{\alpha}_i - \hat{\alpha}_s$$

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

When Friedman (1962) advocated competition in school markets, he had in mind that reputation reflects value added, as noted in point 1 above. The fact that student ability is imperfectly measured implies that reputation also reflects student composition, however. Point 2 thus provides an endogenous desire for better peers. Point 3 also shows that some parents will rationally prefer more “diverse” schools.

4.3. A benchmark public school system. Consider first a public system in which there is no school choice; students are randomly allocated to schools. Hence the mean ability at each school is equal to the population mean: $\hat{\alpha}_s = 0$. The relative diversity is π^{NS} and effort satisfies:

$$(16) \quad d'(e^{NS}) = \pi^{NS} \frac{\gamma_1}{\gamma_0}.$$

Thus, effort is lower than at the efficient allocation, characterized by $d'(e^*) = \frac{\gamma_1}{\gamma_0}$. This reflects the impact of the incentive compatibility constraint.

The government allows a set of schools n to operate. Some are of high and some of low productivity; specifically $n_H + n_L = n > 1$, so there is excess capacity, but $n_H < 1$. Thus, high productivity schools cannot serve all students. This implies that the probability that a student attends a high productivity school is $p_H = \frac{n_H}{n_H + n_L}$. This is lower than the probability at an efficient allocation, where all high productivity schools would be fully utilized.

We now explore whether introducing choice and competition would improve performance along either the student effort or the school value added dimension.

4.4. A public school system with school choice. Consider now a non-selective public system that features school choice. Students can apply to any school and are admitted based only upon their lottery number. To determine the resulting allocation, consider first how students would compare schools with the same value added. If one school happens to have a better peer group, this makes it more attractive, and given the random allocation system, no student loses by applying. If all students apply, then the random admissions rule implies that the school's innate log ability will be $\hat{\alpha}_s = 0$, and its relative diversity $\pi_s = \pi^{NS}$. Hence, in this case it is not possible to have an equilibrium in which one school has a better peer group than another.

Moreover, all else equal, students prefer high value added, hence the high productivity schools will fill to capacity. After these are filled, $1 - n_H$ students will be allocated to low productivity schools (and will be strictly worse off than the rest).²⁹ Thus we may conclude:

Proposition 4. *A non-selective public school system and school choice ensures that high productivity schools are filled to capacity. All schools have the same relative diversity, and so effort is the same as in the benchmark public system.*

Thus, school choice can enhance performance relative to our benchmark public system.³⁰

4.5. A public system including selective schools. Many jurisdictions attempt to improve educational quality by introducing selective public schools. These admit students based only on testing merit and exist in cities throughout the world.³¹ Our model suggests their introduction can *worsen* performance.

To see this, suppose the government allows n_{PS} selective public schools to operate. As in Section 2, the subscript indicates their partially selective nature—admissions depend upon the admissions assessment but not upon background. These schools operate alongside $n_H + n_L$ non-selective schools, where $n_H + n_L + n_{PS} > 1$.³² Let v_{PS} denote the value added at the selective schools (we return to its level below). Additionally, suppose that students wish to go to these schools in equilibrium—we work out sufficient conditions for this to be the case below.

²⁹ Hence the equilibrium allocation has $f(\psi|s) = n_s f(\psi)$ where $\int_0^{n_L} n_s ds = 1 - n_H$.

³⁰ In this type of scenario, the positive impact of choice can also arise from effects on effort. For an illustration, suppose a public system initially had some spatial segregation, *and* there are no transportation costs. School choice would ensure all high productivity schools are utilized, and eliminate cross-school variations in peer quality. This would increase effort in at least some schools. See Angrist, Bettinger, Bloom, Kremer, and King (2002) and Angrist, Dynarski, Kane, Pathak, and Walters (2010) for some evidence that is consistent with the prediction that choice can enhance performance.

³¹ For example, they exist in New York and Paris, and their expansion is under discussion in Lima and Santiago.

³² The last condition is necessary whenever $n_L > 0$; otherwise equilibrium does not exist.

Only students with admissions scores above a threshold—those with $\tau_i \geq \tau^{PS}$ —are admitted, where the cutoff τ^{PS} solves:

$$n_{PS} = 1 - F(\tau^{PS}) \equiv \int_{\tau^{PS}}^{\infty} \int_{-\infty}^{\infty} f(b, \tau) db d\tau,$$

We assume that once admitted to the selective sector students are randomly assigned to its schools.³³ The remaining students, those with $\tau < \tau_{PS}$, apply to the non-selective public schools, where their priority again depends on a randomly assigned number.

This setup implies that the distribution of student characteristics at the selective public schools will be:

$$f^{PS}(\psi) = \begin{cases} \frac{f(\psi)}{1-F(\tau^{PS})}, & \tau_i > \tau^{PS}, \\ 0, & \tau_i \leq \tau^{PS}. \end{cases}$$

Similarly, the distribution will be the same at all the non-selective schools, and given by the population distribution truncated at τ^{PS} :

$$f^{NS*}(\psi) = \begin{cases} \frac{f(\psi)}{F(\tau^{PS})}, & \tau_i \leq \tau^{PS}, \\ 0, & \tau_i > \tau^{PS}. \end{cases}$$

where the asterisk in the superscript differentiates this sector from the case, denoted by NS , in which only non-selective schools exist.

Now consider the impact that introducing partially selective schools has on effort. Let the relative diversity at the selective public schools be given by:³⁴

$$\pi^{PS}(\tau^{PS}) = \pi_s(\Gamma^{PS}),$$

Similarly, let $\pi^{NS*}(\tau^{PS})$ be the relative diversity for the non-selective sector. The selectivity of the elite sector reduces its relative diversity, and thus we have $\pi^{PS}(\tau^{PS}) < \pi^{NS}$. Due to the cream skinning, it is also the case that: $\pi^{NS*}(\tau^{PS}) < \pi^{NS}$. Thus, the introduction of elite schools *reduces* effort at all schools.³⁵

Beyond moving effort further from that in the efficient allocation, the introduction of selective public schools can lower the pressure schools experience to raise their value added.

Proposition 5. *Holding value added in the partially selective public sector, v^{PS} , constant, and assuming that the elite and high productivity non-selective schools cannot serve all students ($n_{PS} + n_H < 1$), then for*

³³ More specifically, once in this segment students can once again apply to all its schools, with priority given by their randomly assigned numbers. We could allow for further stratification by test scores within the selective segment; we abstract from it for simplicity as we deal with such stratification in the next section.

³⁴ Below we discuss how to derive the relative diversity of a school with a selected population.

³⁵ Additionally, $\lim_{\tau_{PS} \rightarrow \infty} \pi^{NS*}(\tau^{PS}) = \pi^{NS}$, and $\lim_{\tau_{PS} \rightarrow \infty} e^{NS*}(\tau^{PS}) = e^{NS}$ and $\lim_{\tau_{PS} \rightarrow \infty} e^{PS}(\tau^{PS}) = e^{FS}$.

selectivity (τ^{PS}) sufficiently high it is case that $u(\psi_i, s^{PS}, \Gamma^{PS}) > u(\psi_i, s^{NS^*}, \Gamma^{PS})$ for any student, where s^{PS} is any partially selective public school, and s^{NS^*} is any non-selective school.

For high enough selectivity, individuals prefer a selective school even if it is of low productivity; indeed, even if $v_{PS} < v_L$. Again, in our setting individuals value reputation, and thus may opt for better peer groups at the expense of lower value added. Put otherwise, given the existence of an imperfection—the fact that the government constrains entry into the school sector—choice and selectivity can worsen system performance. Note also that Proposition 5 provides a first illustration of the *anti-lemons effect*: signaling concerns can allow the entry of a selective school even if it has no productivity advantage.

4.6. Free private entry: The anti-lemons effect. For a final case, consider the benchmark public non-selective scenario, and suppose the government allows free entry by private schools. In particular, public schools can choose to become private (i.e., selective), and collect the voucher \bar{P} . This allows us to suppose that the total number of high and low productivity schools is n_L and n_H respectively, but that their status as public or private is endogenous.

In general, admissions policies can be very complex. We simplify matters by assuming that private schools respond only to students' desire for a high second period wage. We show that this leads to admissions policies such that each school specializes in a single ability level. Next we show that under the appropriate conditions, there is a mixed public-private equilibrium, with high ability students going to the private sector.

4.6.1. Optimal admissions standards. Consider a private school s that admits a range of students: individuals whose expected ability $\hat{\alpha}_i$ is in the interval $[\underline{\alpha}_s, \bar{\alpha}_s]$, with a continuous distribution over this interval given by $h_s(\hat{\alpha})$, with $h_s(\underline{\alpha}) > 0$. Under free entry, another school could try to compete away these students. What would happen if a more selective school—one with a higher $\underline{\alpha}$ —offered admission to qualifying individuals at school s ? Such students' utility at a private school is given by:

$$(17) \quad u(\hat{\alpha}_i, e_s, s) = \gamma_0 \{ \log((1-T)Y_i) - d(e_s) \} + \gamma_1 E \{ \pi_s (\hat{\alpha}_i + e_s + v_s) + (1 - \pi_s) R_s \}.$$

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

Proposition 6. *Suppose that school s admits only students with expected log 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. Thus, under free entry the only possible equilibrium for private schools is to specialize in students of a specific ability, $\hat{\alpha}_s$.

Hence, in equilibrium private schools will choose to be perfectly selective, as we assumed in Section 2.

4.6.2. *The anti-lemons effect.* Starting from our benchmark public equilibrium, suppose the government allows the operation of private schools with admissions criteria $\hat{\alpha}_s$. A student is admitted to a private school s only if her expected innate ability, $\hat{\alpha}_i$, is at least $\hat{\alpha}_s$. Students remaining in the public sector are randomly assigned to schools. Since high productivity private schools can always out compete low productivity schools, as long as the size of the private sector is less than n_H , all private schools are of high productivity.

From Proposition 5 we know that if private entry is permitted, it will take place to some extent. Proposition 6 implies that if a private school with admissions standard α_s enrolls students, all individuals with ability $\hat{\alpha}_i > \alpha_s$ will also choose to attend private schools tailored to their ability. Hence, if there is a mixed private/public equilibrium, it is characterized by a cutoff α^{FS} , such that individuals with $\hat{\alpha}_i > \alpha^{FS}$ attend fully selective (FS) private schools, while those with $\hat{\alpha}_i < \alpha^{FS}$ remain in the non-selective public sector.

To summarize, to some extent fully selective private schools with better reputations (higher valued added and better peer groups) will displace public schools with weaker reputations. The private sector will be stratified, and once a high productivity school catering to individuals of a specific innate ability exists, it will have a stable market position. These outcomes are the *anti-lemons effect*—asymmetric information regarding ability that is signaled via school reputation allows entry by selective schools that cream skim the best students from the non-selective sector.

4.6.3. *The extent of the private sector.* Will the private sector grow until all public schools are displaced? In order to have an equilibrium featuring both types of schools, it must be the case that a student with innate ability α^{FS} (the one at the cutoff) is indifferent between a public and a private school, and that students with lower ability strictly prefer the public sector.

In general, such an equilibrium exists, but characterizing it is complicated because effort in the public and private sectors is not the same. As a result, before presenting the general existence result we discuss a simple case that highlights the role of information. Specifically, suppose log effort is zero ($e = 0$) and all schools are of low productivity ($n_H = 0$). If a mixed equilibrium with cutoff α^{FS} exists, the expected ability of students in the non-selective public sector is $A(\alpha^{FS}) = E\{\alpha | \hat{\alpha} \leq \alpha^{FS}\}$. Since in this case the school system plays a purely informational role, the marginal student will be indifferent between a private and a

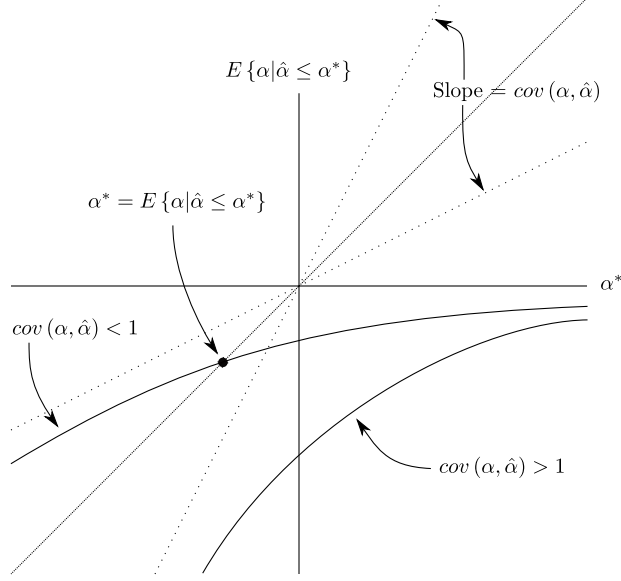


FIGURE 1. Equilibrium Cutoff Ability - α^{FS}

public school if:

$$(18) \quad \alpha^{FS} = A(\alpha^{FS}).$$

At this point, the marginal private entrant no longer offers a better reputation. This is possible if student characteristics ($\psi_i = \{b_i, \tau_i\}$) are not too informative regarding innate ability.

Proposition 7. *Suppose that all students select zero log effort. If the admissions assessment and family background are sufficiently uninformative ($cov(\alpha, \hat{\alpha}) < 1$), there is a unique cutoff α^{FS} solving (18). Conversely, if $cov(\alpha, \hat{\alpha}) > 1$ then there are only fully selective private schools in equilibrium.*

This equilibrium is illustrated in Figure 1. Students with ability $\hat{\alpha}_i \geq \alpha^{FS}$ choose a private school that yields a future expected wage of $\hat{\alpha}_i + v_L$. Students with ability $\hat{\alpha}_i \leq \alpha^{FS}$ remain in the public sector and expect a wage of $A(\alpha^{FS}) + v_L \leq \hat{\alpha}_i + v_L$.

A feature of this equilibrium is that as the covariance between α and $\hat{\alpha}$ falls, the cutoff ability, α^{FS} rises. In the limit, if $\hat{\alpha}$ is completely uninformative, then expected ability is the prior population mean, and we have $A(\alpha) = 0$ for all α , and hence $\alpha^{FS} = 0$; half of the market consists of private schools.

4.6.4. The General Case—Endogenous Effort. More generally, as long as $cov(\alpha, \hat{\alpha}) < 1$ there exists an equilibrium featuring both types of schools. The entrance of selective private schools complicates the characterization of the equilibrium because it affects the relative diversity, and hence effort in both sectors. Specifically, suppose high productivity schools enter the selective private sector first. If the size of this sector

is less than the supply of high productivity schools (namely $1 > n_H > 1 - F(\alpha^{FS})$), then the remainder go into the non-selective public sector. Given an equilibrium cutoff α^{FS} , the fraction of high productivity schools in the public sector is:

$$\lambda_H(\alpha^{FS}) = \max\{0, n_H - (1 - F(\alpha^{FS}))\}.$$

Given the random allocation of students to public schools, the expected value added in this sector as a (continuous) function of the equilibrium cutoff α^{FS} is:

$$(19) \quad v^{NS*}(\alpha^{FS}) = v_H \lambda_H(\alpha^{FS}) + v_L (1 - \lambda_H(\alpha^{FS})).$$

When $v^{NS*}(\alpha^{FS}) = v_L$, all high productivity schools are in the private sector, and the margin between the two sectors features a low productivity school. When this is not the case, the marginal student selects between a high productivity private school and a public school with expected value added: $v_H > v^{NS*}(\alpha^{FS}) > v_L$.

Next, consider the effect private entry has on the effort incentive faced by students in the public sector:

$$(20) \quad d'(e^{NS*}) = \frac{\gamma_1}{\gamma_0} E \left\{ \frac{\partial w^{NS*}(\alpha^{FS}, t_i)}{\partial t_i} | \psi_i \right\}.$$

The term in brackets (the effect of the graduation assessment t upon future wages) can be bounded, leading to the following proposition.

Proposition 8. *Consider students in public schools when these co-exist with private schools as defined by a cutoff α^{FS} . The marginal effect of effort upon future expected wages is increasing with α^{FS} and satisfies:*

$$\frac{\rho^t}{\rho + \rho^t} \geq E \left\{ \frac{\partial w^{NS*}(t, \alpha^{FS})}{\partial e} | \psi_i \right\} \geq \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t}, \quad \text{and}$$

$$\begin{aligned} \lim_{\alpha^{FS} \rightarrow \infty} E \left\{ \frac{\partial w^{NS*}(t, \alpha^{FS})}{\partial e} | \psi_i \right\} &= \frac{\rho^t}{\rho + \rho^t} \\ \lim_{\alpha^{FS} \rightarrow -\infty} E \left\{ \frac{\partial w^{NS*}(t, \alpha^{FS})}{\partial e} | \psi_i \right\} &= \frac{\rho^t}{\rho + \rho^b + \rho^\tau + \rho^t}. \end{aligned}$$

Thus, when the non-selective public sector is large the incentive for effort among public school students is close to that observed under a pure public system (Proposition 2). As the selective private sector grows, the incentive for effort falls to that observed in a private school.

This result also shows that effort is bounded above and below. This fact combined with the observation that student utility at a selective private school grows without bound with selectivity, $\hat{\alpha}_s$, implies the existence of a mixed public-private equilibrium when the covariance between ability and $\hat{\alpha}$ is less than one.

Proposition 9. *Suppose that the admissions assessment and family background are sufficiently uninformative ($\text{cov}(\alpha, \hat{\alpha}) < 1$). There is an equilibrium characterized by α^{FS} such that students with ability $\hat{\alpha}_i < \alpha^{FS}$ attend non-selective public schools, while those with $\hat{\alpha}_i > \alpha^{FS}$ attend fully selective private schools. Students at public schools exert higher effort, and this effort falls with the size of the private sector.*

A feature of this equilibrium is that holding the productivity of the private schools fixed, raising that of the public sector increases the size of the latter. Increasing the precision of the graduation assessment raises effort in both sectors. Hence it unambiguously enhances system performance, though the impact upon the size of the two sectors is ambiguous. Finally, increasing the precision of the admissions assessment has an ambiguous effect. At the margin it reduces effort in the selective sector and hence reduces its performance, and hence its size. However, if $\text{cov}(\alpha, \hat{\alpha})$ is large, it can lead to all schools choosing to be selective, reducing overall system performance.

In summary, a concern for school reputation in a model with school choice leads to an endogenous preference for better peers. If left unchecked, this leads to an equilibrium that stratifies students by ability, and choose lower effort than at the efficient of benchmark public equilibrium. None of the systems we studied achieves the first best. However, the most efficient equilibrium combines free entry with lottery-based admissions. This ensures that the high value added schools are fully utilized, while mitigating the adverse incentives created by the anti-lemons effect. This case corresponds roughly to the current rules governing the entry charter schools under the No Child Left Behind legislation in the United States.

5. CONCLUDING DISCUSSION

In this paper, we study a competitive educational market in which schools are able to acquire a reputation for quality, as measured by the achievement of their graduates. Building upon Holmstrom's (1999) model, we show that schools can enhance their reputations in two ways. First, they can increase the quality of instruction. This effect captures Friedman (1962)'s intuition that competition for a good reputation can lead to gains in school productivity.

Second, schools can improve their reputations by attracting better students. The first order consequence of this is an *anti-lemons* effect. Akerlof (1970) showed that asymmetric information in a competitive market may result in the exit of high quality sellers. In contrast, reputational concerns may lead to entry by relatively

small schools that serve students within a specific ability range. This results in stratification, with the most able students attending the schools with the best reputations and subsequently earning the highest incomes, while the least able remain in the worst schools.³⁶

The anti-lemons effect implies that competition for a good reputation need not lead to gains in educational performance. Holmstrom (1999) shows that equilibrium effort decreases with the precision of expected ability.³⁷ Competition leads to greater school selectivity and a more precise estimate of ability; as a result it produces lower *relative diversity*—the weight the market places upon individual-specific measures of skill as opposed to school reputation. This in turn leads students to exert lower academic effort. The anti-lemons effect can also lead parents to prefer more selective schools even if these do not offer higher value added (from either better instruction or from causal human capital externalities).

As we discussed, our model can reconcile some empirical findings that are not consistent with models that rely upon peer effects. Moreover, it rationalizes some variation in preferences over relative diversity; holding mean reputation fixed, more able students prefer less selective schools, while the converse is true for less able students.

Finally, observe that signaling via school reputation has implications that differ from Spence’s (1973) celebrated signalling model. In his model, signals are valuable because they are expensive to acquire, and the cost of acquiring a “good” signal falls with ability. This leads high ability individuals to invest in signals that might otherwise have no value. In our setting, his model predicts that students’ preference for selective schools would lead them to engage in unproductive activities that increase the likelihood of admission to such schools. For instance, buying tutoring that teaches to the test, or using advice books to better game the admissions process.³⁸ This prediction reinforces our result more egalitarian admissions policies, combined with better measures of individual performance, as has been recommended by Bishop (1997), may enhance the performance of the market for education.

This also suggests that future research should explore models that include both reputation and Spence-type signaling effects. Such models would have a rich set of implications that may help us understand the complex structure of modern education markets.

³⁶ 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.

³⁷ Gibbons and Murphy (1992) find that this can explain the greater reliance upon performance pay for older CEOs.

³⁸ On amazon.com the keywords “school admissions” leads to 2,197 entries with titles like “Testing for Kindergarten: Simple Strategies to Help Your Child Ace the Tests for: Public School Placement, Private School Admissions, Gifted Program Qualification”.

REFERENCES

- AKERLOF, G. A. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84(3), 488–500.
- ALMOND, D., AND J. M. CURRIE (2011): "Human Capital Development Before Five," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and C. D., vol. 4. Elsevier.
- 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.
- ANGRIST, J. D., S. M. DYNARSKI, T. J. KANE, P. A. PATHAK, AND C. R. WALTERS (2010): "Who Benefits from KIPP?," Discussion Paper 15740, NBER, Cambridge, MA.
- 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 E. ROUSE (2009): "School vouchers and student achievement: Recent evidence and remaining questions," *Annual Review of Economics*, January(17-42).
- 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.
- BISHOP, J. H. (1997): "The Effect of National Standards and Curriculum-Based Exams on Achievement," *The American Economic Review*, 87(2), 260–264.
- BLACK, S. (1999): "Do better schools matter? Parental valuation of elementary education," *The Quarterly Journal of Economics*, 114(2), 577–599.
- 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*.
- 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.
- CUNHA, F., J. HECKMAN, AND S. SCHENNACH (2010): "Estimating the Technology of Cognitive and Non-Cognitive Skill Formation," Discussion Paper 15664, NBER, Cambridge, MA 02138.
- 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.
- DEGROOT, M. H. (1972): *Optimal Statistical Decisions*. McGraw-Hill Book C., New York, NY.
- 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.

- DUFLO, E., P. DUPAS, AND M. KREMER (2008): "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya," Discussion Paper 14475, NBER.
- 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.
- GIBBONS, R., AND K. J. MURPHY (1992): "Optimal Incentive Contracts in the Presence of Career Concerns," *Journal of Political Economy*, 100, 30–52.
- GROGGER, J. (2009): "Speech Patterns and Racial Wage Inequality," Discussion Paper 08.13, University of Chicago, Harris School Working Paper Series 08.13.
- HARRIS, M., AND B. HOLMSTRÖM (1982): "A Theory of Wage Dynamics," *Review of Economic Studies*, 49(3), 315–333.
- HASTINGS, J., AND J. WEINSTEIN (2008): "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of Economics*, 123(4), 1373–1414.
- 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.
- 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): "School markets: The impact of information approximating schools' effectiveness," Mimeo, Columbia University.
- NEAL, D. (2009): "Private schools in education markets," in *Handbook of research on school choice*, ed. by M. Berends, M. Springer, D. Balou, and H. Walberg. Routledge, New York, NY.
- NECHYBA, T. J. (2006): "Income and Peer Quality Sorting in Public and Private Schools," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, and F. Welch, vol. 2. Elsevier.

- 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.
- PATHAK, P. A. (2011): “The Mechanism Design Approach to Student Assignment,” *Annals Review of Economics*, 3.
- POP-ELECHES, C., AND M. URQUIOLA (2011): “The consequences of going to a better school,” Mimeo, Columbia University.
- 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.
- OLON, G. (1992): “Intergenerational Income Mobility in the United-states,” *American Economic Review*, 82(3), 393–408.
- SPENCE, M. (1973): “Job Market Signaling,” *The Quarterly Journal of Economics*, 87(3), 355–374.
- VALENZUELA, J., C. BELLEI, AND D. DE LOS RIOS (2010): “Segregacion escolar en Chile,” in *Fin de Ciclo?*, ed. by S. Martinic, and G. Elacqua. UNESCO, Santiago.
- 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

A.1. Proof of Proposition 1: Proposition 1 follows from: i) our assumptions regarding the distribution of innate ability, and ii) equations (1), (2), and (3) for public, selective public, and private schools, respectively. \square

A.2. Proof of Proposition 2: Consider the public schools first. The market estimates log skill, θ_i , using two signals, t_i and R_{s_i} . The estimate given the first signal is $E\{\theta_i|R_{s_i}\} = \hat{\alpha}_s^{NS} + e_s + v_s$, with precision ρ . Additionally, t is a direct measure of skill, with precision ρ^t . Hence, Bayesian learning implies the first two lines in Table 2. The remaining cases are similar, and follow from the fact that for selective public schools we have $E\{\theta_i|R_{s_i}\} = \hat{\alpha}_s^{PS} + e_s + v_s$ with precision $\rho^\tau + \rho$; for private schools we have $E\{\theta_i|R_{s_i}\} = \hat{\alpha}_s^{FS} + e_s + v_s$ with precision $\rho^b + \rho^\tau + \rho$. \square

A.3. Proof of Proposition 3: At the optimum both budget constraints bind. Hence, once in the labor market, individuals consume:

$$C^1(\psi_i, t_i) = A_i(\psi_i, t_i) E(\psi_i) V(\psi_i).$$

Note that $A_i(\psi_i, t_i) = E\{\exp(\alpha_i)|\psi_i, t_i\} = \exp\left(E\{\alpha_i|\psi_i, t_i\} + \frac{\text{var}\{\alpha_i|\psi_i, t_i\}}{2}\right)$. Hence, taking logs:

$$\log(C^1(\psi_i, t_i)) = E\{\alpha_i|\psi_i, t_i\} + \frac{\text{var}\{\alpha_i|\psi_i, t_i\}}{2} + e(\psi_i) + v(\psi_i).$$

One can substitute this, along with (10) into the objective function to get:

$$\begin{aligned} E\{u(C^0(\psi_i), C^1(\psi_i, t_i), E(\psi_i))|\psi_i\} &= E\{\gamma_0(\log(C_0(\psi_i) - d((E(\psi_i)))) + \gamma_1 \log(C_1(\psi_i, t_i))|\psi_i\} \\ &= \gamma_0(\log(Y_i - K(v(\psi_i))) - d(e(\psi_i))) \\ &\quad + \gamma_1(E\{\alpha_i|\psi_i, t_i\} + \text{var}\{\alpha_i|\psi_i, t_i\} + e(\psi_i) + v(\psi_i)). \end{aligned}$$

Taking the first order conditions of this expression yields the conditions in Proposition 3. Notice that since $K'(v)$ is increasing and unbounded, this ensures a unique solution to:

$$K'(v_i^*) = \frac{\gamma_1}{\gamma_0}(Y_i - K_H(v_i^*)).$$

Finally, this implies:

$$\frac{\partial v_i^*}{\partial Y_i} = \frac{1}{(K''(v_i^*)\gamma_0/\gamma_1 + K'(v_i^*))} > 0$$

and hence value added increases with income. \square

A.4. Proof of Proposition 4. Students can apply to as many schools as they wish, and they are given priority according to a randomly generated number. Since effort is a second order effect, small differences in reputation or relative diversity will cause a large number of students to change their application choices. Hence, the only equilibrium is for all schools to have the same reputation and relative diversity; i.e., the distribution of students at each school is the same as the population distribution.

Students strictly prefer high productivity schools, and given that schools with the same characteristics in equilibrium have the same distribution of types, then those with high productivity are chosen first. Once these are filled, the students will randomly sample from the remaining low productivity schools.

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

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

Thus (recall that we omit the effect of the variance term when taking the expectation of wages—this has no effect upon the qualitative results), the payoff of a student with characteristics ψ_i who enters school s is:

$$\begin{aligned} u(c^0(\psi_i), c^1(\psi_i, t_i), e(\psi_i)) &= \gamma_0 \{ \log((1 - T) \bar{Y}) - d(E(\psi_i)) \} \\ &+ \gamma_1 E \{ \pi^{NS} (t_i - \hat{e}_s - v_s (T \bar{Y})) + \hat{e}_s + v_s (T \bar{Y}) | \psi_i \}. \end{aligned}$$

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

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

as in the benchmark public system. \square

A.5. Proof of Proposition 5. The payoff from attending a selective school is:

$$(21) \quad u(\psi_i, s \in PS, \Gamma^{PS}) = \gamma_0 \{ \log((1 - T) Y) - d(e_s) \} + \gamma_1 \{ \pi^{PS} (\Gamma^{PS}) (\hat{\alpha}_i + e_s + v_s) + (1 - \pi^{PS} (\Gamma^{PS})) R^{PS} (\tau^{PS}) \}.$$

Observe that the relative diversity π^{PS} falls with τ^{PS} , but is bounded below (this is shown formally in Proposition 8). This implies that log effort is bounded below. Moreover, the expected ability of students at a selective school rises without bound as τ^{PS} rises. Together these imply that school reputation rises without bound with τ^{PS} . Moreover, the utility of a student with ability τ^{PS} rises with approximately a slope of 1. However, for schools in the non-selective sector, as τ^{PS} rises, the payoff approaches that of the non-selective public school equilibrium. For a student with ability τ^{PS} in the non-selective sector, utility rises with her expected graduation assessment, t_i , with a slope of approximately π^{NS} ; hence utility in the selective sector rises more quickly than in the non-selective sector with τ^{PS} . This implies that regardless of its value added, if τ^{PS} is sufficiently large all students will wish to attend the selective sector. \square

A.6. Proof of Proposition 6. Observe that $(\alpha, \hat{\alpha}_i, t'_i)$ has a multivariate normal distribution with strictly positive covariance matrix and zero means. Let $f(\alpha, \hat{\alpha}_i, t'_i)$ be the corresponding probability density function, and $f(\hat{\alpha}_i, t'_i)$ the density when α is integrated out. This implies:

$$E \{ E \{ \alpha | \hat{\alpha}_i \in [\underline{\alpha}_s, \bar{\alpha}_s], t'_i \} | \hat{\alpha}_i \} = \frac{\int_{-\infty}^{\infty} \hat{\alpha} (t_i, s) f(\hat{\alpha}_i, t'_i) dt'_i}{\int_{-\infty}^{\infty} f(\hat{\alpha}_i, t'_i) dt'_i}$$

where:

$$\hat{\alpha}(s_i, t_i) = \left\{ \frac{\int_{\underline{\alpha}_s}^{\bar{\alpha}_s} \int_{-\infty}^{\infty} \alpha f(\alpha, \hat{\alpha}, t'_i) h(\hat{\alpha}) d\alpha d\hat{\alpha}}{\int_{\underline{\alpha}_s}^{\bar{\alpha}_s} \int_{-\infty}^{\infty} f(\alpha, \hat{\alpha}, t'_i) h(\hat{\alpha}) d\alpha d\hat{\alpha}} \right\}$$

is the expected ability of an individual who attended school s and has graduation assessment score t_i . When computing her future wage, the student has to take into account the correlation between $\hat{\alpha}_i$ and t_i . However, to work out the effect of increasing selectivity upon wages, it is sufficient to work out the effect upon expected ability conditional upon the recommendation t_i if this always has the same sign.

We have:

$$\begin{aligned}\frac{\partial \hat{\alpha}(t_i, s)}{\partial \underline{\alpha}_s} &\geq \frac{f_s(\underline{\alpha}_s, t'_i)}{f_s(t'_i)} \left\{ -\frac{\int_{-\infty}^{\infty} \alpha f(\alpha, \underline{\alpha}_s, t'_i) h(\underline{\alpha}_s) d\alpha}{f_s(\underline{\alpha}_s, t'_i)} + \hat{\alpha}(t_i, s) \right\} \\ &\geq \frac{f_s(\underline{\alpha}_s, t'_i)}{f_s(t'_i)} \{ \hat{\alpha}(t_i, s) - E\{\alpha_i | t_i, \hat{\alpha}_i = \underline{\alpha}_s\} \} > 0.\end{aligned}$$

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 α_i in the sense of Milgrom (1981). \square

A.7. Proof of Proposition 7. As stated, the expected innate ability of a student in a non-selective public school given that a fully selective private sector with cutoff α^{FS} exists is $A(\alpha^{FS}) = E\{\alpha | \hat{\alpha} \leq \alpha^{FS}\}$. This expected ability can be computed using results on conditional expectations of normal random variables with truncation. From Birnbaum (1950) we have:

$$E\{X | Z \geq z\} = \mu R(z),$$

where X and Z are standard normal random variables with variances σ_X^2 and σ_Z^2 , respectively; $\mu = E\{XZ\}$, $R(z) = \frac{f(z)}{1-F(z)}$ 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 Z have normal distributions, then:

$$(22) \quad E\{X | Z \leq z\} = E\{X\} - \frac{\text{cov}\{X, Z\}}{\sigma_Z} R\left(\frac{E\{Z\} - z}{\sigma_Z}\right).$$

To use these facts, note that the log innate ability in the public sector is characterized by a normal distribution truncated at α^{FS} ; hence $A(\alpha^{FS}) = E\{\alpha | \hat{\alpha} \leq \alpha^{FS}\}$. From Proposition 1, $E\{\alpha\} = E\{\hat{\alpha}\} = 0$, and $\text{var}(\hat{\alpha}) = 1/(\rho^\tau + \rho^b + \rho)$. Notice that $\text{var}(\alpha) = \text{cov}\{\alpha, \tau\} = 1/\rho$, and $\text{cov}\{\alpha, \zeta b\} = \text{cov}\{\zeta b + \epsilon^b, \zeta b\} = \zeta^2/\rho^y$, which implies the covariance:

$$(23) \quad \text{cov}(\alpha, \hat{\alpha}) = \frac{\rho^\tau/\rho + \zeta^2\rho^b/\rho^y}{(\rho^\tau + \rho^b + \rho)}.$$

Applying (22) implies:

$$(24) \quad E\{\alpha | \hat{\alpha} \leq \alpha^{FS}\} = -\frac{\text{cov}(\alpha, \hat{\alpha})}{\sqrt{\rho^\tau + \rho^b + \rho}} R\left(-\bar{\alpha}\sqrt{\rho^\tau + \rho^b + \rho}\right).$$

Moreover, from Birnbaum (1950) we have that

$$(25) \quad x + \frac{1}{x} \geq R(x) \geq x$$

and for $x \geq 0$, and

$$(26) \quad \lim_{x \rightarrow -\infty} R(x) = 0.$$

Thus for $\alpha^{FS} \leq 0$:

$$\text{cov}(\alpha, \hat{\alpha}) \alpha^{FS} \geq E\{\alpha | \hat{\alpha} \leq \alpha^{FS}\} \geq \text{cov}(\alpha, \hat{\alpha}) \alpha^{FS} + \frac{\text{cov}(\alpha, \hat{\alpha})}{\alpha^{FS}(\rho^\tau + \rho^b + \rho)}.$$

Note that from (24) $A(\alpha^{FS})$ is a continuous, increasing function. Moreover, the positive correlation between α and $\hat{\alpha}$ implies $A(\alpha^{FS}) < 0$, and hence if a solution α^{FS} to (18) exists, it must be negative. Let $h(\alpha^{FS}) = A(\alpha^{FS}) - \alpha^{FS}$, and note that $h(0) < 0$. We have:

$$\begin{aligned}h(\alpha^{FS}) &\geq \text{cov}(\alpha, \hat{\alpha}) \alpha^{FS} + \frac{\text{cov}(\alpha, \hat{\alpha})}{\alpha^{FS}(\rho^\tau + \rho^b + \rho)} - \alpha^{FS} \\ &\geq (\text{cov}(\alpha, \hat{\alpha}) - 1) \alpha^{FS} + \frac{\text{cov}(\alpha, \hat{\alpha})}{\alpha^{FS}(\rho^\tau + \rho^b + \rho)}.\end{aligned}$$

Hence $\text{cov}(\alpha, \hat{\alpha}) < 1$ and the right hand side goes to ∞ as $\alpha^{FS} \rightarrow -\infty$, and by the intermediate value theorem we have an α^{FS} such that $h(\alpha^{FS}) = 0$. Conversely, for $\alpha^{FS} < 0$ and $\text{cov}(\alpha, \hat{\alpha}) > 1$ we have:

$$\alpha^{FS} > \text{cov}(\alpha, \hat{\alpha}) \alpha^{FS} \geq E\{\alpha | \hat{\alpha} \leq \alpha^{FS}\},$$

and hence a solution cannot exist. \square

A.8. Proof of Proposition 8. Let \bar{s} denote a school in the public sector, which is characterized by $\hat{\alpha} \leq \alpha^{FS}$. Let $w(t, \alpha^{FS})$ be the expected wage of an individual from the public sector who enters the labor market with a graduation assessment score t . Since $\frac{\partial t_{i\bar{s}}}{\partial e_{i\bar{s}}} = 1$, the level of effort, $e_{\bar{s}}(\hat{\alpha}, \alpha^{FS})$, in this sector solves:

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

We assume that in equilibrium value added and effort 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 (22) to compute $E\{\alpha | t^\alpha, \hat{\alpha} \leq \alpha^{FS}\}$ by letting $X(t^\alpha) = E\{\alpha | t^\alpha, \hat{\alpha}\}$ and $Z = \hat{\alpha}$, and take all expectations conditional upon t^α . We have:

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

Using the Bayesian learning rule:

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

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

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

We also have:

$$\begin{aligned} E\{Z(t^\alpha)\} &= \frac{\rho^t}{\rho + \rho^t} t^\alpha, \\ \text{var}\{Z(t^\alpha)\} &= \text{var}(\alpha | t^\alpha) + \text{var}(\hat{\epsilon}), \\ &= \frac{1}{\rho + \rho^t} + \frac{1}{\hat{\rho}}, \\ &= \frac{\rho + \hat{\rho} + \rho^t}{\hat{\rho}(\rho + \rho^t)} \\ &= \sigma_Z^2 \end{aligned}$$

Finally the covariance:

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

We now use formula (22) to obtain:

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

We can now determine the marginal effect of effort upon future wages:

$$(27) \quad \frac{\partial w(t, \alpha^{FS})}{\partial e} = \frac{\rho^t}{\rho + \rho^t} \left(1 - \frac{\hat{\rho}}{(\rho + \hat{\rho} + \rho^t)} R' \left(\frac{\frac{\rho^t}{\rho + \rho^t} t^\alpha - \alpha^{FS}}{\sigma_Z} \right) \right).$$

From the inequalities (25) on R we have $1 \geq R' \geq 0$, from which we conclude:

$$\frac{\rho^t}{\rho + \rho^t} \geq \frac{\partial w(t, \alpha^{FS})}{\partial e} \geq \frac{\rho^t}{\rho + \hat{\rho} + \rho^t},$$

and therefore:

$$\frac{\rho^t}{\rho + \rho^t} \geq E \left\{ \frac{\partial w(t, \alpha^{FS})}{\partial e} | \hat{\alpha} \right\} \geq \frac{\rho^t}{\rho + \hat{\rho} + \rho^t},$$

Finally, from the properties (25) and (26) of R it is also the case that:

$$(28) \quad \lim_{\hat{\alpha} \rightarrow \infty} E \left\{ \frac{\partial w(t, \alpha^{FS})}{\partial e} | \hat{\alpha} \right\} = \frac{\rho^t}{\rho + \rho^t}.$$

$$(29) \quad \lim_{\hat{\alpha} \rightarrow -\infty} E \left\{ \frac{\partial w(t, \alpha^{FS})}{\partial e} | \hat{\alpha} \right\} = \frac{\rho^t}{\rho + \hat{\rho} + \rho^t}.$$

A.9. Proof for Proposition 9. The payoff to a student with ability α^{FS} who attends a non-selective public, high productivity selective private, and low productivity selective private school are:

$$\begin{aligned} u^{NS}(\alpha^{FS}) &= \gamma_0 \{ \log((1-T)Y) - d(e(\alpha^{FS})) \} + \gamma_1 \{ A(\alpha^{FS}) + e(\alpha^{FS}) + v^{NS}(\alpha^{FS}) \}, \\ u_H^{FS}(\alpha^{FS}) &= \gamma_0 \{ \log((1-T)Y) - d(e(\alpha^{FS})) \} + \gamma_1 \{ \alpha^{FS} + e^S + v_H \}, \text{ and} \\ u_L^{FS}(\alpha^{FS}) &= \gamma_0 \{ \log((1-T)Y) - d(e(\alpha^{FS})) \} + \gamma_1 \{ \alpha^{FS} + e^S + v_L \}, \end{aligned}$$

respectively. Since $A(\alpha^{FS})$ is bounded above, for large α^{FS} we have $u_H^{FS}(\alpha^{FS}), u_L^{FS}(\alpha^{FS}) > u^{NS}(\alpha^{FS})$. For small α^{FS} we have $A(\alpha^{FS}) \simeq \text{cov}(\alpha, \alpha^{FS}) \alpha^{FS}$. This, combined with the fact that the other terms are bounded and $\text{cov}(\alpha, \alpha^{FS}) < 1$ implies $u_H^{FS}(\alpha^{FS}), u_L^{FS}(\alpha^{FS}) < u^{NS}(\alpha^{FS})$. Hence there exists $\alpha_L^{FS} > \alpha_H^{FS}$ satisfying:

$$\begin{aligned} u_H^S(\alpha_H^{FS}) &= u^{NS}(\alpha_H^{FS}) \\ u_L^S(\alpha_L^{FS}) &= u^{NS}(\alpha_L^{FS}). \end{aligned}$$

The equilibrium depends upon the number of high productivity schools. There are three cases to consider. If $n_H > 1 - F(\alpha_H^{FS}) > 1 - F(\alpha_L^{FS})$, then the number of high productivity schools is greater than the demand for selective schools at $\alpha^{FS} = \alpha_H^{FS}$. Hence this defines an equilibrium at which the marginal student is indifferent between a high productivity private school and the non-selective sector.

If $n_H < 1 - F(\alpha_L^{FS})$, the number of high quality school is not sufficient to meet the demand, and hence the marginal student is indifferent between the non-selective sector and a low productivity private school. Finally, if $1 - F(\alpha_H^{FS}) \geq n_H \geq 1 - F(\alpha_L^{FS})$ then we let α^{FS} satisfy $n_H = 1 - F(\alpha^{FS})$. At this value, all selective private schools are high productivity, and hence the marginal student is choosing between a high productivity selective private school and a school in the selective sector, and will choose the selective sector. This is an equilibrium, because in order for the selective sector to increase in size it must use a low productivity school. However, since $\alpha^{FS} < \alpha_L^{FS}$ then $u_L^S(\alpha^{FS}) < u^{NS}(\alpha^{FS})$, and students would not choose to attend a low productivity selective private school. This demonstrates the existence of an equilibrium. The last result follows from the properties of relative diversity for the selective and non-selective sectors as given in the previous proposition.