High Scores but Low Skills

Liqun Liu and William S. Neilson*

March 2009

Abstract

In this paper college admissions are based on test scores and students can exert two types of effort: real learning and exam preparation. The former improves skills but the latter is more effective in raising test scores. In this setting the students with the lowest skills are no longer the ones with the lowest aptitude, but instead are the ones closest to the borderline for college admission. Increased access to college leads to greater income inequality between college graduates and non-graduates. Overall, the ability to study for the test leads to higher expected test scores but lower skills.

Keywords: college admission; standardized test; tournament; income inequality; teaching to the test
JEL Classification: I21, I28, J24

* We want to thank Scott Carrell, David Figlio, Scott Gilpatric, and Philip Trostel for insightful discussions and Amy Hopson for excellent research assistance.
Liu: Private Enterprise Research Center, Texas A&M University
Neilson: Department of Economics, University of Tennessee, Knoxville

"The tutors may not actually help you speak better English, but your scores will certainly be better." – an anonymous student

"In my heart, I don't agree with these tutoring practices, but when everyone else takes their classes and your children don't, you are afraid they won't be as competitive." – an anonymous parent

1. Introduction

Every year, graduating high school students compete with each other for a limited number of college freshman seats. The reason for such competition is the so-called "college premium" that can be broadly defined to include not only the wage premium for college graduates compared to high school graduates, but also the job satisfaction and social status associated with a college degree. One channel through which pre-college students compete is testing. Because college applicants' skills are not perfectly observable at the time of admission decisions, colleges that want to admit students with high skills rely on various standardized tests to assess applicants' skill levels. In the extreme case, admission is solely based on each student's scores on a college-entrance exam.[1]

College entrance examinations, or standardized tests in general, are designed in such a way that, other things being equal, students with higher skills would score higher on such tests. Therefore, it is somewhat surprising to hear college professors and employers of high school graduates complain about a situation in which the average test score gets higher from one cohort to another, but the average level of skills actually becomes lower. This phenomenon, dubbed "high scores but low skills," is widely

---

[1] For example, a dominant annual national college entrance exam exists in Mainland China, Hong Kong, Taiwan, and South Korea. For an introduction to the college entrance examination system of China, see Wang (2006). On the other hand, American colleges have adopted more comprehensive admission policies in which standardized test scores are of relatively moderate importance.

observed, but is especially prevalent in cultures that place particularly high values on education and in countries that have admission policies solely based on college entrance examination scores.[2] Then why do standardized tests designed to test a student's skills generate scores that can misrepresent his/her skill level? How can we explain the phenomenon of "high scores but low skills?"

It has been widely recognized that tests conducted in a classroom environment are limited in assessing test-takers' skills that are useful for dealing with real life challenges. For example, standardized tests are typically designed to assess test-takers' cognitive skills rather than non-cognitive skills, such as emotional and social skills, that are also critical for employment or further education. Within the category of cognitive skills, tests tend to focus on knowledge that is well established and skills that can be easily measured or graded. As a result, controversial knowledge and skills requiring creativity are usually omitted in tests. But being creative and being able to process competing and controversial information is important for effective working and learning.[3]

Therefore, standardized tests by their very nature can only test a subset of useful skills. As a result, while a well-balanced effort at improving overall skills helps raise the test score, it is usually not as effective as a more focused approach: teaching/studying to the test. There is evidence that drilling students on content known to be on a test can

---

[2] Complaints about high scores but low skills are much more widely spread in China than in the US, most likely because in China test scores play a much more important role in college admission. In particular, Li (2005) observed that, "Tests are an important and necessary means of evaluating students, but the problem with examination-oriented education is that both teaching and learning focus on dealing with exams, whereas education in morality and aesthetics – as well as students' physical and mental health – are neglected, resulting in uneven student development." (pp. 308-309)

[3] In a survey article by Hanushek (1986), it is pointed out that almost all studies of earnings that include both quantity of schooling and achievement (test score) measures find significant effects of quantity that are independent of achievement differences, suggesting that the education process may have multiple outputs, some of which are very poorly measured by test scores (achievements).

significantly improve the students' performance on the test, but the resulting high test scores do not necessarily translate into corresponding knowledge and skills.[4]

In this paper, we distinguish between two kinds of effort by a pre-college student in her preparation for application to college: the effort on real learning focusing on skill improvement, and the effort on exam preparation focusing on raising test scores. The latter effort is more effective in raising test scores but less efficient in improving skill levels than the former. For example, to prepare for the verbal part of the SAT, one can either read a selection of classic novels which include many of the literary words emphasized in the SAT or directly memorize a list of SAT words. While the former strategy exposes students to some of the vocabulary, it also provides lessons about the human condition and introduces the reader to different cultures, places, and times. On the other hand, the latter strategy is much more effective in preparing for the specific exam.

In a tournament model of college admissions with these two types of effort, we predict a very robust phenomenon of high scores but low skills.[5] More precisely, we compare a benchmark case in which only learning is possible to a setting in which both types of effort are possible, and find that the ability to prepare for the college entrance examination leads to higher scores but lower skills. It also leads to other behavioral changes. When test preparation is not possible increases in the college premium lead to higher skills and higher test scores, but when test preparation is possible increases in the

---

[4] See Smith and Fey (2000) and Burger and Krueger (2003).

[5] While these theoretical predictions are consistent with some casual observations of employers and college professors, the direct empirical evidence of high scores but low skills is still lacking. Further complications arise from changes in the SAT test. Still, from 1994, when major changes to SAT were made, to 2005, when another round of major changes were made a year later in 2006, the average SAT-V score for college-bound seniors has increased from 499 to 508, and the average SAT-M score has increased from 504 to 520 (College Board SAT 2007). At the same time access to college increased, with 12.4 million students enrolled in 1994 and 14.2 million enrolled in 2005 (US Census Bureau, 2006).

college premium lead to higher scores but lower skills.[6] The reason is that higher college premiums increase the payoffs from studying for the college entrance examination, so students substitute away from learning and toward test preparation. The ability to prepare for the college entrance examination does not change who gets into college, though, with higher aptitude students having higher expected test scores and greater chances of admission in both cases.[7]

Our model also generates some other, surprising findings. One concerns the relationship between aptitude and skills. When test preparation is not possible there is a direct relationship between aptitude and skills: higher aptitude students learn more and acquire more skills. In particular, the students with the lowest skills are the ones with the lowest aptitude. When test preparation is possible, however, this direct relationship fails, and now the students with the lowest skills are the ones who are nearest the cutoff for getting into college. The reason is that these borderline students have the most at stake from test preparation so they concentrate their efforts there, while those with the lowest aptitude have little chance of getting into college and so concentrate on acquiring skills.

A second surprising implication of the model concerns access to college. Standard analysis suggests that increasing access to higher education would reduce

---

[6] Katz and Murphy (1992) document the increase in the college premium over time. Contributing to the increased demands for skilled labor are skill-biased technological change (Autor, Katz, and Krueger, 1998) and the expansion of international trade (Wood, 1998).

[7] The positive relationship between test scores and aptitude found in this paper is consistent with the positive relationship between SAT scores and freshman grade point averages found by some SAT validity studies (e.g. Bridgeman et al., 2000, and Camara and Echternacht, 2000). This is because the "high scores but low skills" phenomenon is about variations between cohorts, whereas the finding of a positive correlation between SAT scores and college performance is mostly about variations within cohorts. Much of the predictive power assigned to the SAT for within-cohort comparisons comes from the unobserved student talents or omitted demographic and socioeconomic characteristics (Rothstein, 2004). In other words, it is the difference in talent within a cohort that contributes to the positive relationship between scores and skills. As we show in this paper, on the other hand, the negative relationship between scores and skills across cohorts is caused by the changing environment of college entrance competition, in which the average talent (or the talent distribution) of a cohort remains constant over time.

income inequality between college graduates and non-college graduates, as the changes in supply reduce the wages of graduates and increase the wages of non-graduates. When test preparation is not possible, a second mechanism works in the same direction, with increased access enhancing the expected marginal payoff (in terms of expected admission to college) from skill acquisition for low aptitude students but decreasing the expected marginal payoff for high aptitude students, thereby reducing the productivity of college graduates and increasing the productivity of non-graduates. When test preparation is possible, though, this second mechanism gets reversed. The increased access causes low aptitude students to substitute toward test preparation and away from learning, so that they acquire fewer skills. At the same time, increased access makes high aptitude students surer of admission, freeing them to substitute toward learning and away from test preparation, increasing their skills. Consequently, when students can prepare explicitly for college entrance examinations, increased access to higher education can increase, rather than reduce, income inequality.

The work in this paper is complementary to the work of Epple, Romano, and Seig (2002, 2003, 2006, 2008). In these papers students differ in income and ability, and ability is measured by a test score. Our model separates ability from test scores, and derives as a result that higher ability students earn higher test scores, on average. The model of Epple et al. treats the distribution of test scores as exogenous where we treat it as endogenous, determined by ability and two types of effort. On the other hand, we consider a single college and treat its enrollment capacity as exogenous, while Epple et al. consider a hierarchy of colleges competing for the best students. Such a hierarchy is compelling, but well beyond the scope of this paper, and a melding of the two models is worthy of further research.[8]

Section 2 introduces the college admission tournament model and notation. Section 3 explores how expected test scores and skill levels vary with aptitude, and Section 4 explores what happens when colleges lower the admission standard. Section 5 establishes the high scores but low skills phenomenon, showing that when the admission standard is held constant, the ability to substitute test preparation for real learning leads students at every aptitude level to have higher expected test scores but fewer skills. Section 6 offers concluding remarks, and all proofs are collected in an appendix.

2. College Admission Tournaments

A society contains a unit mass of students indexed by their aptitude $a > 0$. Individuals can expend costly effort on two activities, and the effectiveness of these activities is determined by aptitude. One of the activities is effort devoted to real learning, denoted e_L, and it impacts the individual's pre-college productivity. In particular, if an individual with aptitude a exerts e_L units of effort on real learning, her pre-college productivity is $f(ae_L)$, where $f' > 0$ and $f'' < 0$. In the remainder of the paper we refer to the value of $f(ae_L)$ as *skills*. The other activity is explicit preparation for the college entrance exam, and this effort is denoted e_E. Unlike real learning, test preparation

---

[8] Besides the issues related to peer effects as modeled by Epple et al., different schools may generate different wage premia. For example, Hoekstra (2008) finds that wage premia are higher for students who attend a state's flagship university.

has no direct impact on skills.[9] The total cost of effort is given by the function $c(e_E + e_L)$, with $c' > 0$ and $c'' > 0$.

The reason for exerting effort toward test preparation is that it improves the chances of getting into college. College, in turn, provides a fixed increment π to productivity, so that an individual with skill level y has post-college productivity $y + \pi$, where π is assumed to be net of college costs and independent of y.[10] An individual who does not attend college experiences no gain in productivity, so pre-college and post-college productivity levels are the same. Labor markets are competitive and workers are paid according to their productivity.

College admissions are based solely on entrance exam scores. In particular, the college sets a threshold score s, and any student who scores s or higher gains admittance and realizes the college premium. A student's test score t is a random variable influenced by her aptitude a, effort e_L on learning, and effort e_E on test preparation, and given by

$$t = a(\theta e_E + e_L) + \varepsilon,$$

---

[9] In that e_E is less efficient than e_L in building skills but more effective than e_L in raising test scores, the effort on exam preparation here is similar to "teaching to the test". In the US, there have been heated media discussions about how "teaching to the test", in response to mandatory state testing that relates student performance to school funding and teacher pay raise, has severely narrowed the curriculum, inflating scores at the cost of in-depth classroom instruction. Underscoring the strong incentives high-stakes state testing provides for schools to alter the ways in which educational services are delivered, Figlio (2006) found evidence that schools assign harsher punishments to low-performing students than to high-performing students during the testing window in order to reshape the testing pool.

[10] The assumption of the college premium being independent of the pre-college skill level is based on the following considerations. First, studies on changes in college premiums suggest that the college premium is mainly affected by technology progress and international trade at the macro level (see footnote 6), which has little to do with pre-college efforts at the individual level. Second, there exists some direct evidence that college premiums are unaffected by pre-college skill variations. For example, Grogger and Eide (1995) found evidence that skills attained prior to college, as measured by standardized test scores and high school grades, had no effect on the change in the college wage premium for men. On the other hand, changes in the college major distribution explained a significant portion of the rise in the male college wage premium. As for what determine the college premium or rate of return to schooling in general, two theories – human capital and signaling – have been advanced in the literature (see, for example, Groot and Oosterbeek 1994 and Weiss 1995). More recently, Fang (2006) finds that human capital enhancement accounts for at least two-thirds of the college wage premium, with the remainder arising from ability signaling.

where θ is a parameter governing the relative effectiveness of the two types of effort for raising the test score, and ε is a mean-zero random variable with distribution function $G(\cdot)$ and density function $g(\cdot)$.[11] Note that both types of effort impact test scores, and that aptitude magnifies the impact of effort in the same way as it does in the production function.[12] Because the effort on test preparation does not contribute to productivity, for a student to exert any test preparation effort it must be the case that $\theta > 1$ so that direct test preparation has a greater effect on test scores than general (but productivity-enhancing) learning.[13]

The student is admitted to college if $t \geq s$, which occurs if

$$\varepsilon \geq s - a(\theta e_E + e_L).$$

Consequently, the probability of acceptance is

$$1 - G(s - a(\theta e_E + e_L)).$$

Besides having a mean of zero, the density function g is assumed to be bell-shaped, so that $g'(\varepsilon) > 0$ when $\varepsilon < 0$, $g'(\varepsilon) < 0$ when $\varepsilon > 0$, and $g'(0) = 0$.

The objective of a student with aptitude a is to choose e_L and e_E to maximize

---

[11] Epple, Romano, Sarpça, and Sieg (2006) also assume that colleges observe ability imperfectly, but in the context of a negotiation model where colleges compete for students instead of a tournament model in which students compete for admission.

[12] Both kinds of efforts are mainly made by pre-college students, but they also include any teachers' and parents' efforts. Focusing on the effects of efforts (rather than the effects of monetary inputs) on skill building and test scores is consistent with the empirical findings that spending per student is not significantly correlated to test scores, dropout rates or continuation rates to higher levels of schooling, or to labor market outcomes (Hanushek 1986). It is also consistent with the empirical finding that working during school has a quantitatively large negative impact on academic performance (Stinebrickner and Stinebrickner 2003).

[13] We are not aware of any estimates of θ for SAT or for any other test. Existing studies on the effects of coaching on SAT scores generally cast doubt on some commercial coaching programs' claim of being able to boost the combined verbal and math score by 100 to 140 points. For example, Powers and Rock (1999) found the coaching effect estimates for SAT I verbal (math) scores to range from 6 to 12 points (13 to 26 points); the effects of coaching found in Briggs (2004) are from 3 to 20 points for SAT-V and from 10 to 28 points for SAT-M. However, these coaching effect estimates do not directly lead to estimates for θ, because we do not know how many hours are behind the coaching examined in those coaching effect studies and how large a gain in scores would be made by "real" learning of the same amount of time.

$$f(ae_L) + \pi[1 - G(s - a(\theta e_E + e_L))] - c(e_E + e_L). \tag{1}$$

The first-order conditions are

$$af'(ae_L) + \pi a g(s - a(\theta e_E + e_L)) - c'(e_E + e_L) = 0 \tag{2a}$$

$$\theta\pi a g(s - a(\theta e_E + e_L)) - c'(e_E + e_L) = 0. \tag{2b}$$

The solutions to the first-order conditions (2a) and (2b) can be written $e_L(a,s,\pi,\theta)$ and $e_E(a,s,\pi,\theta)$, and the resulting expected test score is denoted

$$T(a,s,\pi,\theta) = a(\theta e_E(a,s,\pi,\theta) + e_L(a,s,\pi,\theta)).$$

The second-order sufficient condition for a maximum is satisfied if the matrix M of second partials is negative definite, where

$$M = \begin{pmatrix} a^2 f'' - \pi a^2 g' - c'' & -\theta\pi a^2 g' - c'' \\ -\theta\pi a^2 g' - c'' & -\theta^2\pi a^2 g' - c'' \end{pmatrix}. \tag{3}$$

The second-order condition implies that

$$a^2 f'' - \pi a^2 g' - c'' < 0 \tag{4a}$$

$$D = a^2\pi g'[(\theta - 1)^2 c'' - a^2\theta^2 f''] - a^2 f'' c'' > 0. \tag{4b}$$

We assume throughout the paper that expressions (4a) and (4b) hold.

Combining the first-order conditions (2a) – (2b) yields the following:

$$f'(ae_L) = (\theta - 1)\pi g(s - a(\theta e_E + e_L)). \tag{5}$$

This equation implies that if the individual's maximization problem possesses an interior solution, so that the chosen levels of e_L and e_E are both strictly positive, it must be the case that $\theta > 1$. The reasoning is straightforward. An increase in e_L impacts both the student's skill level and her probability of admission to college, while an increase in e_E affects only the probability of admission. If $\theta \leq 1$, then e_E is less effective than e_L for improving test scores and the student would never devote effort to test preparation. The only reason that test preparation is attractive is because it is more effective than learning for improving entrance exam scores.

The model implicitly assumes that students are "standard-takers," that is, no individual student can influence the college's choice of standard *s* by changing effort. This allows us to avoid explicitly modeling students' responses to other students' effort decisions, which would greatly complicate the analysis.[14] It is also a good approximation of reality, as many colleges receive thousands of applications each year. The college has a fixed number of slots it can fill and sets the standard *s* so that, in expectation, it fills all of the slots. For the students, though, this standard can be considered exogenous.

The distribution of aptitude is given by the function $H(a)$, with density function $h(a)$. Given the standard *s*, the other parameters of the model (π and θ), and the two distribution functions (G and H), the mass of students admitted to college is given by

$$A(s) = \int [1 - G(s - T(a, s, \pi, \theta)] h(a) da \, .$$

Colleges are capacity constrained and can accommodate a fraction *r* of the population. As a consequence, the standard *s* satisfies

$$A(s) = r.$$

3. Aptitude, College Admissions, and Skill Acquisition

We begin analysis of the college admissions problem by describing behavior across aptitude levels. This is a reasonable place to begin, because students are identical

---

[14] The standard tournament theory of Lazear and Rosen (1981) would have students choose effort levels that are best responses to the equilibrium effort levels of all other students. The standard-taking assumption allows us to consider students individually.

except for aptitude, and aptitude impacts the effectiveness of both the effort for skill generation and the effort for exam preparation.

College admission is determined solely by test scores, and the first proposition characterizes the relationship between aptitude and expected test scores (all proofs are collected in the appendix).

Proposition 1. *Expected test scores increase with aptitude.*

Since students are indexed by their aptitude, Proposition 1 identifies who is expected to go to college. According to the model expected test scores are governed in part by aptitude, but also by effort on learning and exam preparation. The proposition shows that, in equilibrium, exam preparation does not distort the relationship between aptitude and test scores, and the highest aptitude students are the ones expected to get into college.

Proposition 1 adds some structure to equilibrium. In particular, the fact that expected test scores increase with aptitude means that those above some threshold aptitude level $\bar{a}$ expect to be admitted and those below the threshold aptitude level expect to be denied admission, and $\bar{a}$ solves

$$T(\bar{a}, s, \pi, \theta) = s .$$

Since $s - T(a,s,\pi,\theta)$ is decreasing in a and $g(s - T(a,s,\pi,\theta))$ is bell-shaped, we have $g' > 0$ for students with aptitude $a > \bar{a}$ and $g' < 0$ for students with aptitude $a < \bar{a}$.

Given the nature of the model, with the college premium being the same for everyone who goes to college regardless of aptitude or skills, the question of who goes to college in and of itself has no welfare implications. The real welfare question has to do with skills. The ability to prepare for the tests leads students to devote effort to test preparation, but in equilibrium has no impact on who gets into college, and so the test preparation effort is socially wasteful. If students substitute test preparation for learning, skill acquisition is impacted. The next proposition describes this impact across aptitude levels.

Proposition 2. *Skills increase with aptitude for students who expect to be admitted to college* (*i.e.* $a > \overline{a}$) *and decrease with aptitude for students who expect to be denied admission* (*i.e.* $a < \overline{a}$).

Proposition 2 establishes that skill levels are a U-shaped function of aptitude, which is both surprising and troubling, particularly because it implies that the students with the lowest skills are the ones who are closest to the borderline for admission to college. The reason is clear. For these students studying for the test has the largest impact on whether they get into college and earn the college premium, and so they study the most for the test and exert the least effort toward learning. Since learning impacts productivity but test preparation does not, these borderline students have the lowest skills. In contrast, in a world without test preparation, the students with the lowest skills would be the ones with the lowest aptitude.[15]

Propositions 1 and 2 together imply that one gets very different pictures of the college population depending on whether one looks at skills or test scores. Of course, in expectation college-bound students have higher test scores than non-college-bound ones, and Proposition 1 confirms that these are also the students with the highest aptitude. So,

---

[15] See Lemma 1 in the appendix.

in expectation the worst college-bound student has higher aptitude than the best non-college-bound student. On the other hand, Proposition 2 says that the least-skilled college bound student has the same skills as the least-skilled non-college-bound student. The most-skilled college-bound student, which is the student with the highest aptitude, may have more skills than the most-skilled non-college-bound student, which is the student with the lowest aptitude, so the average skill level of college-bound students may exceed the average skill level of non-college-bound ones. The two skill ranges overlap, however, and from a skill perspective the college-bound may not look much different from the non-college bound. This would not be the case if the entrance examination were not coachable, because then both skills and expected test scores increase with aptitude, and the college-bound outperform the non-college bound on both measures of performance.[16]

4. Increased Access to College

In a standard supply and demand analysis, increasing access to college increases the supply of college graduates but lowers the supply of high school graduates. Holding demand in the two markets constant, the increased access should lead to a lower wage for graduates and a higher wage for non-graduates, reducing inequality. Essentially, the inequality reduction comes through a reduced college premium as the increased number of college graduates compete it away.

Supply and demand analysis requires an endogenous college premium, which we do not have in this model. Nevertheless the same result, that increased college access reduces the wage disparity between high school and college graduates, would also obtain

---

[16] See Lemma 1 in the appendix.

in our model if test preparation were prohibited.[17] When college becomes more accessible those who were unlikely to be admitted before now find their chances of admission improved, and the payoff from increasing test scores rises. The only way to raise test scores when test preparation effort is impossible is through learning, which also raises skills, and so the increased accessibility makes lower aptitude students learn more and higher aptitude students learn less. This makes the skill levels more equal. And, unlike the supply and demand analysis, the inequality reduction here comes from more equal skills, not from a smaller college premium.

The next proposition explores the relationship between college accessibility and skill inequality when students can exert exam-specific effort.

Proposition 3. *When college becomes more accessible skill levels rise for students who expect to be admitted to college* (*i.e.* $a > \overline{a}$) *and fall for students who expect to be denied admission* (*i.e.* $a < \overline{a}$).

When it is possible to prepare for the test, increased accessibility to college leads to increased inequality in skills. The mechanism is as follows. When college is inaccessible, low-aptitude students are unlikely to get in, and so do not devote much effort to exam preparation. Instead, they devote it to learning, which increases their productivity. At the same time, high-aptitude students are working hard to make sure they do get in, and so they have high test preparation effort and relatively low learning, which makes their productivity lower. But, when college becomes more accessible, those low-aptitude students suddenly have a better shot at getting into college and so move

---

[17] See Lemma 1 in the appendix.

more of their effort into exam preparation. This lowers their productivity when they do not get into college. On the other hand, those high-aptitude students do not need to worry as much about getting in, and so they can devote less effort to test preparation and more to learning, and their productivity levels rise.

5. High Scores but Low Skills

So far we have looked at who, as measured by aptitude, gets into college and how efforts to get into college impact skill acquisition for different aptitude levels. We have not, so far, explored whether efforts to get into college, in the form of preparing specifically for the entrance examination, lead to higher or lower skill levels overall. Instead, Propositions 1 and 2 only tell us that test scores rise with aptitude but skill levels do not, with the lowest skill levels coming from those closest to the college admissions threshold. The purpose of this section is to determine whether students participate in more or less real learning when they are able to prepare specifically for the entrance examination.

An intermediate result concerns the impact of the college premium, and it is a partial equilibrium result rather than a general equilibrium result:

Proposition 4. *Hold the admissions standard constant. Then when the college premium* π *increases, exam-specific effort increases, expected test scores increase, and real learning decreases.*

Proposition 4 states that increases in the college premium make test scores rise but skill levels fall, in line with the high scores but low skills phenomenon. The result is a partial equilibrium result because of the assumption that the admissions standard does not rise when the college premium goes up. If the standard stays the same but test scores rise, more students are admitted to college and any college capacity constraint is violated. Combining Proposition 4 with Proposition 3 suggests that when the college premium rises and the required test score to get into college adjusts upward, the impact on skills is negative for high-aptitude students and ambiguous for low-aptitude students. Everyone learns less because of the wage premium effect which entices them to shift effort away from learning and toward exam preparation, and the higher admissions standard exacerbates this effect for high-aptitude students as they fight for their seats at college. The total effect is ambiguous for low-aptitude students, though, because while the higher wage premium makes exam preparation more attractive, college is harder to get into making learning more attractive.

The real issue, though, is not whether an increase in the college premium exacerbates the high scores but low skills problem, but whether the ability to improve test scores through exam-specific effort leads to the problem in the first place. The next proposition establishes the high scores but low skills phenomenon, again in a partial equilibrium framework.

Proposition 5. *Assume that the standard is held constant. For any individual, the existence of exam-specific effort leads to lower skills but higher expected test scores.*

The proposition establishes that, fixing the admissions standard and any aptitude level, the individual exerts less effort toward learning but nevertheless earns higher test scores when test preparation is possible than when it is not. So, it is the ability to devote effort to test preparation that leads to the high scores but low skills phenomenon.

6. Conclusion

Because colleges with limited spots want to admit students with high skills but cannot observe applicants' skills at the time of the admission decision, they have to rely on a college entrance exam or a standardized test to screen applicants. In this paper, we have demonstrated that an admission policy based solely on the score of a college entrance examination would lead to a "high scores but low skills" phenomenon: compared to a scenario in which the entrance examinations are not coachable, the ability to study explicitly for the college entrance examination leads to higher test scores but lower skills at all aptitude levels. This reduction in skills occurs because students substitute effort away from real learning and toward exam preparation, and this takes place even though the ability to prepare for the exam has no impact on who gets into college, in equilibrium.

The across-the-board reduction in skill levels is troubling, if only because it reduces a country's productivity and competitiveness. The ability to prepare for entrance exams has other worrisome implications, as well. First, it changes the monotone relationship between aptitude and skills. Without exam preparation higher aptitude students have higher skills, but with exam preparation the lowest skills are obtained by students close to the borderline for college admission. The intuition behind this result is straightforward: borderline students have the most to gain from exam preparation, and so they do the most of it, neglecting real learning in the process. This U-shaped skills/aptitude pattern leads to an enhanced “high scores but low skills” effect among students actually enrolled in college, because the students who were barely admitted are among the least skilled in the entire applicant pool.

Another worrisome implication concerns the ability of an increase in access to higher education to equalize incomes across the population. When students cannot prepare for the examination, increasing access leads to more skill acquisition among low-aptitude students and less skill acquisition among high-aptitude students, leading to less inequality between college graduates and non-graduates. When students can prepare for the college entrance examination, on the other hand, this pattern is reversed, and increased access leads to greater income inequality.

The findings of this paper suggest that a welfare improvement, caused by an across-the-board increase in skill levels, could be obtained if college admission policies were changed so that exam preparation were impossible, or at least less effective in raising scores. It should be noted, though, that while the analysis in the paper only concerns test preparation, it could be extended to other factors that improve chances of admission with only a tenuous impact on skills, such as overzealous participation in extracurricular activities. However, it is in no way clear how to make college admissions less dependent on an entrance examination or a standardized test. For example, GPAs are not comparable across schools, and focusing on them may cause grade inflation. Essays and recommendation letters are subjective and are costly to evaluate. More importantly, allowing less objective criteria than test scores may induce rent-seeking behavior by admission officers and could be the grounds for corruption. For those countries that rely solely on college entrance examination scores to assign college seats, they may well have adopted the fairest and the most efficient college admission policy, given the prevailing culture and institutions, even though the high scores but low skills problem is an inevitable consequence of such a policy.[18]

---

[18] Li (2005) observed that, given the limited resources of higher education in China, “in comparison with other selection methods, tests are still the fairest and the most sensible.” (p360)

APPENDIX

Lemma 1. *When exam preparation is impossible (so that e_E is restricted to zero), higher aptitude students have both higher expected test scores and more skills. Furthermore, lowering the admission standard leads to more learning and higher skills for low aptitude students and less learning and lower skills for high aptitude students, and increases in the college premium lead to increases in learning.*

Proof. Rewrite the objective function in (1) to preclude the possibility of test preparation effort:

$$f(ae_L) + \pi[1 - G(s - ae_L))] - c(e_L). \tag{A1}$$

The first-order condition is

$$af'(ae_L) + \pi ag(s - ae_L) - c'(e_L) = 0, \tag{A2}$$

and the second-order condition is

$$a^2 f'' - \pi a^2 g' - c'' < 0$$

which is guaranteed by (4a). Let e_L^* denote the solution to (A2), and let T^* denote the corresponding expected test score given by $T^* = ae_L^*$. Both are functions of a, s, and π. We get

$$\frac{dT^*}{da} = \frac{af' + c''e_L^* + \pi ag}{c'' + \pi a^2 g' - a^2 f''} > 0\,,$$

which establishes that expected test scores increase with aptitude. Also,

$$\begin{aligned}\frac{df(ae_L^*)}{da} &= e_L^* f' + a\frac{de_L^*}{da} f' \\ &= \frac{af' + c''e_L^* + \pi ag}{c'' + \pi a^2 g' - a^2 f''} f' > 0,\end{aligned}$$

which implies that higher aptitude students obtain higher skills.

One can compute

$$\frac{de_L^*}{ds} = \frac{\pi a g'}{c'' + \pi a^2 g' - a^2 f''}. \tag{A3}$$

This derivative has the same sign as g', so is positive when $a > \bar{a}$ and negative when $a < \bar{a}$. Consequently, reducing the standard leads to more learning and higher skills for low-aptitude students and less learning and lower skills for high-aptitude students.

Finally,

$$\frac{de_L^*}{d\pi} = \frac{ag}{c'' + \pi a^2 g' - a^2 f''} \geq 0,$$

so increasing the college premium leads to more learning and higher skills. ■

Proposition 1. *Expected test scores increase with aptitude.*

Proof. Differentiating the first-order conditions with respect to a yields, in matrix form,

$$M\begin{pmatrix} de_L / da \\ de_E / da \end{pmatrix} = \begin{pmatrix} -f' - ae_L f'' - \pi g + (\theta e_E + e_L)\pi a g' \\ -\theta \pi g + (\theta e_E + e_L)\theta \pi a g' \end{pmatrix},$$

where the matrix M is given in (3). By expression (5) we have $f' = (\theta - 1)\pi g$, and substituting this into the above expression yields

$$M\begin{pmatrix} de_L / da \\ de_E / da \end{pmatrix} = \begin{pmatrix} -(\theta - 1)\pi g - ae_L f'' - \pi g + (\theta e_E + e_L)\pi a g' \\ -\theta \pi g + (\theta e_E + e_L)\theta \pi a g' \end{pmatrix}.$$

Solving and simplifying yields

$$\frac{de_L}{da} = \frac{1}{D}\left[ae_L f''(\theta^2 \pi a^2 g' + c'') + (\theta e_E + e_L)(\theta - 1)\pi a g' c'' + (\theta - 1)\theta^2 \pi^2 a^2 g' g\right]$$

and

$$\frac{de_E}{da} = \frac{1}{D}\left[-ae_L f''(\theta\pi a^2 g' + c'') + (\theta e_E + e_L)\pi a g'(\theta a^2 f'' - (\theta - 1)c'') - \theta\pi a^2 g(f'' + (\theta - 1)\pi g')\right]$$

Since the expected test score is

$$T = a(\theta e_E + e_L),$$

we have

$$\frac{dT}{da} = a\left(\theta\frac{de_E}{da} + \frac{de_L}{da}\right) + \left(\theta e_E + e_L\right)$$

which simplifies to

$$\frac{dT}{da} = -\frac{a^2\theta f''}{D}\left[c''(e_L + e_E) + \theta\pi a g\right] > 0\,. \qquad ■$$

Proposition 2. *Skills increase with aptitude for students who expect to be admitted to college (i.e.* $a > \overline{a}$ *) and decrease with aptitude for students who expect to be denied admission (i.e.* $a < \overline{a}$ *).*

Proof. Using the formulas generated in the proof of Proposition 1, we have

$$\begin{aligned}\frac{df(ae_L)}{da} &= e_L f' + a\frac{de_L}{da}f' \\ &= \frac{\theta(\theta - 1)\pi a^2 g' f'}{D}\left[(e_E + e_L)c'' + \theta\pi a g\right]\end{aligned}$$

which has the same sign as g'. Consequently, the derivative is positive when $g' > 0$, which occurs when $a > \overline{a}$, and the derivative is negative when $g' < 0$, which occurs for students with aptitude $a < \overline{a}$. ■

Proposition 3. *When college becomes more accessible skill levels rise for students who expect to be admitted to college (i.e. $a > \overline{a}$) and fall for students who expect to be denied admission (i.e. $a < \overline{a}$).*

Proof. Using the same comparative statics approach as in the other proofs, we can find

$$\frac{de_L}{ds} = -\frac{\pi a g'}{D}(\theta - 1)c'',$$

$$\frac{de_E}{ds} = \frac{\pi a g'}{D}\left[(\theta - 1)c'' - \theta a^2 f''\right],$$

and

$$\begin{aligned} \frac{dT}{ds} &= a\theta\frac{de_E}{ds} + a\frac{de_L}{ds} \\ &= \frac{\pi a^2 g'}{D}\left[(\theta-1)^2 c'' - \theta^2 a^2 f''\right] \qquad \text{(A4)} \\ &= \frac{\pi a^2 g'\left[(\theta-1)^2 c'' - \theta^2 a^2 f''\right]}{\pi a^2 g'\left[(\theta-1)^2 c'' - \theta^2 a^2 f''\right] - a^2 f'' c''} < 1. \end{aligned}$$

College becomes more accessible when r increases. To assess the impact on the admission standard s of an increase in r, differentiate the expression $A(s) = r$ at the end of Section 2 with respect to r, which yields

$$-\int g(s - T(a,s,\pi,\theta))\left[\frac{ds}{dr} - \frac{dT(a,s,\pi,\theta)}{ds}\frac{ds}{dr}\right]h(a)da = 1.$$

Rearranging yields

$$\frac{ds}{dr} = \frac{1}{\int g(s - T(a,s,\pi,\theta))\left[dT/ds - 1\right]h(a)da},$$

which has the same sign as $(dT/ds - 1)$.

By the last line of (A4) we have d$T(a,s,\pi,\theta)$/ds < 1 for all a. Therefore, s decreases when r increases. The impact of this decrease on skills is determined by

$$\begin{aligned}\frac{df(ae_L)}{ds} &= a\frac{de_L}{ds}f' \\ &= -\frac{\pi a^2 g'}{D}(\theta-1)c''f'.\end{aligned}$$

The sign of the derivative is the opposite of the sign of g', and so the effect of increased access has the same sign as g'. We know that $g' > 0$ when $a > \bar{a}$ and $g' < 0$ when $a < \bar{a}$. Hence, skills increase with access when $a > \bar{a}$ and decrease with access when $a < \bar{a}$. ■

Proposition 4. *Hold the admissions standard constant. When the college premium* π *increases, exam-specific effort increases, expected test scores increase, and real learning decreases.*

Proof. The comparative statics derivatives $de_L/d\pi$ and $de_E/d\pi$ are given by the system

$$M\begin{pmatrix} de_L/d\pi \\ de_E/d\pi \end{pmatrix} = \begin{pmatrix} -ag \\ -\theta ag \end{pmatrix}, \tag{A5}$$

where the matrix M is given by (3). Solving yields

$$\frac{de_L}{d\pi} = -\frac{1}{D}(\theta-1)agc'' < 0$$

and

$$\frac{de_E}{d\pi} = \frac{ag}{D}\left[(\theta-1)c'' - \theta a^2 f''\right] > 0.$$

As for the expected test score,

$$\frac{dT}{d\pi} = \theta a \frac{de_E}{d\pi} + a \frac{de_L}{d\pi}$$
$$= \frac{a^2 g}{D}\left[(\theta-1)^2 c'' - \theta^2 a^2 f''\right] > 0. \qquad \blacksquare$$

Proposition 5. *Assume that the standard is held constant. For any individual, the existence of exam-specific effort leads to lower skills but higher expected test scores.*

Proof. Let e_L^*, e_E^*, and T^* be as defined in the proof of Lemma 1, and these values pertain to the case in which exam-specific effort is prohibited. Write e_L, and e_L^* as functions of π. When $\pi = 0$, exam-specific learning serves no purpose and $e_E = 0$. The choice of e_L is then the same as it would be if exam-specific effort was prohibited, so $e_L(0) = e_L^*(0)$. By Proposition 4 and Lemma 1, $de_L(\pi)/d\pi < 0 < de_L^*(\pi)/d\pi$ for all $\pi \geq 0$. Then

$$e_L(\pi) - e_L^*(\pi) = \int_0^{\pi}\left[\frac{de_L(z)}{dz} - \frac{de_L^*(z)}{dz}\right]dz < 0$$

for all $\pi > 0$. It follows that $f(ae_L) \leq f(ae_L^*)$, and skills are lower when exam-specific effort is allowed.

Similarly, let $T(\pi)$ and $T^*(\pi)$ be the expected test scores in when exam-specific effort is allowed and prohibited, respectively. Then $T(0) = T^*(0)$ since $e_E = 0$ and $e_L = e_L^*$ when $\pi = 0$. Furthermore, de_L/dx and de_E/dx are given by the system (A5), and solving for $de_E/d\pi$ yields

$$\frac{de_E}{d\pi} = \frac{1}{D}\left[-ag(c'' + \theta\pi a^2 g') + \theta ag(c'' + \pi a^2 g' - a^2 f'')\right],$$

which can be rearranged to generate

$$\frac{\theta a g}{D}=\left[\frac{1}{c''+\pi a^2 g'-a^2 f''}\right]\cdot\frac{de_E}{d\pi}+\frac{ag}{D}\cdot\left[\frac{c''+\pi a^2 g'}{c''+\pi a^2 g'-a^2 f''}\right].$$

Solving (A5) for $de_L/d\pi$ yields

$$\begin{aligned}\frac{de_L}{d\pi}&=\frac{ag}{D}(c''+\theta^2\pi a^2 g')-\frac{\theta a g}{D}(c''+\theta\pi a^2 g')\\&=\frac{ag}{D}\left[(c''+\theta^2\pi a^2 g')-\frac{(c''+\theta\pi a^2 g')^2}{c''+\theta\pi a^2 g'-a^2 f''}\right]-\left[\frac{c''+\theta\pi a^2 g'}{c''+\theta\pi a^2 g'-a^2 f''}\right]\cdot\frac{de_E}{d\pi}\\&=\frac{ag}{D}\left[\frac{D}{c''+\theta\pi a^2 g'-a^2 f''}\right]-\left[\frac{c''+\theta\pi a^2 g'}{c''+\theta\pi a^2 g'-a^2 f''}\right]\cdot\frac{de_E}{d\pi}\\&=\frac{de_L^{\,*}}{d\pi}-\left[\frac{c''+\theta\pi a^2 g'}{c''+\theta\pi a^2 g'-a^2 f''}\right]\cdot\frac{de_E}{d\pi}.\end{aligned}$$

The derivative with respect to test scores can be written

$$\begin{aligned}\frac{dT}{d\pi}&=a\left[\theta-\frac{c''+\theta\pi a^2 g'}{c''+\pi a^2 g'-a^2 f''}\right]\frac{de_E}{d\pi}+a\frac{de_L^{\,*}}{d\pi}\\&=a\frac{(\theta-1)c''-\theta a^2 f''}{c''+\pi a^2 g'-a^2 f''}\cdot\frac{de_E}{d\pi}+a\frac{de_L^{\,*}}{d\pi}\\&>a\frac{de_L^{\,*}}{d\pi}=\frac{dT^*}{d\pi}.\end{aligned}$$

Thus

$$T(\pi)-T^*(\pi)=\int_0^{\pi}\left[\frac{dT(z)}{dz}-\frac{dT^*(z)}{dz}\right]dz>0$$

for all $\pi > 0$, and test scores are higher when exam-specific effort is allowed. ■

References

Autor, David H., Katz, Lawrence F. and Krueger, Alan B. “Computing Inequality: Have Computers Changed the Labor Market?” *Quarterly Journal of Economics*, 1998, 113 (4), pp. 1169-1213.

Bridgeman, B., McCamley-Jenkins, L., Ervin, N., “Predictions of Freshman Grade-point Average from the Revised and Recentered SAT I: Reasoning Test,” *Research Report 2000-1*, College Entrance Examination Board, 2000, New York.

Briggs, Derek C. “Evaluating SAT Coaching: Gains, Effects and Self-Selection,” in Rebecca Zwick, ed., *Rethinking the SAT: The Future of Standardized Testing in University Admissions*, New York, NY: Taylor and Francis Books, Inc. (RoutledgeFalmer), 2004, pp. 217-233.

Burger, John M., and Krueger, Monte, “A Balanced Approach to High-stakes Achievement Testing: An Analysis of the Literature with Policy Implications,” *International Electronic Journal for Leadership in Learning*, 2003, 7(4). http://www.ucalgary.ca/~iejll

Camara, W.J., and Echternacht, G., “The SAT I and High School Grades: Utility in Predicting Success in College.” *Research Report RN-10,* College Entrance Examination Board, 2000, New York.

College Board SAT, “2007 College-Bound Seniors Total Group Profile Report,” College Board, 2007.

Epple, Dennis, Richard Romano, Sinan Sarpça, and Holger Sieg. “Profiling in Bargaining over College Tuition,” *Economic Journal*, 2006, 116(515), pp. F459-F479.

Epple, Dennis, Richard Romano, and Holger Sieg. "On the Demographic Composition of Colleges and Universities in Market Equilibrium," *American Economic Review*, 2002, 92(2), pp. 310-314.

Epple, Dennis, Richard Romano, and Holger Sieg. "Peer Effects, Financial Aid and Selection of Students into Colleges and Universities: An Empirical Analysis," *Journal of Applied Econometrics*, 2003, 18, pp. 503-525.

Epple, Dennis, Richard Romano, and Holger Sieg. "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education," *Econometrica*, 2006, 74(4), pp. 885-928.

Epple, Dennis, Richard Romano, and Holger Sieg. "Diversity and Affirmative Action in Higher Education," *Journal of Public Economic Theory*, 2008, 10(4), pp. 475-501.

Fang, Hanming. "Disentangling the College Wage Premium: Estimating a Model With Endogenous Education Choices," *International Economic Review*, 2006, 47(4), pp. 1151-1185.

Figlio, David N. "Testing, Crime and Punishment," *Journal of Public Economics*, 2006, 90, pp. 837-51

Frank, Robert H. "The Frame of Reference as a Public Good," *Economic Journal*, 1997, 107, pp. 1832-1847.

Grogger, Jeff and Eric Eide. "Changes in College Skills and the Rise in the College Wage Premium," *Journal of Human Resources*, 1995, 30, pp. 280-310.

Groot, Wim and Hessel Oosterbeek, "Earnings Effects of Different Components of Schooling: Human Capital versus Screening," *Review of Economics and Statistics*, 1994, 76(2), pp. 317-21.

Hanushek, Eric A. "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature*, 1986, 24, pp. 1141-1177.

Hoekstra, Mark L. "The Effect of Attending the State Flagship University on Earnings: A Discontinuity-Based Approach," *Review of Economics and Statistics*, forthcoming.

Katz, Lawrence F. and Murphy, Kevin M. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 1992, 107 (1), pp. 35-78.

Lazear, Edward P., and Rosen, Sherwin. "Rank-Order Tournaments as Optimal Labor Contracts," *Journal of Political Economy*, 1981, 89, pp. 841-64.

Li, Lanqing. *Education for 1.3 Billion*, 2005, Foreign Language Teaching and Research Press, China, and Pearson Education Asia Ltd.

Powers, Donald E., and Donald A. Rock, "Effects of Coaching on SAT I: Reasoning Test Scores," *Journal of Educational Measurement*, 1999, 36(2), pp. 93-118.

Rothstein, Jesse, "College Performance Predictions and the SAT," *Journal of Econometrics*, 2004, 121, pp. 297-317.

Smith, M.L., and Fey, P., "Validity and Accountability of High-stakes Testing," *Journal of Teacher Education*, 2000, 51(5), 334-344.

Stinebrickner, Ralph, and Todd Stinebrickner, "Working during School and Academic Performance," *Journal of Labor Economics* 21 (2003) (2), 473-491.

U.S. Census Bureau, Current Population Survey, October 2006.

Wang, Xiang Bo. “An Introduction to the System and Culture of the College Entrance Examination of China.” *Research Notes RN-28*, College Entrance Examination Board, 2006, New York.

Weiss, Andrew, “Human Capital vs. Signaling Explanations of Wages,” *Journal of Economic Perspective*, 1995, 9(4), pp. 133-154.

Wood, Adrian. “Globalisation and the Rise in Labour Market Inequalities,” *Economic Journal*, 1998, 108, pp. 1463-82.