Gaming a Selective Admissions System*

Frances Xu Lee
Loyola University Chicago

Wing Suen
University of Hong Kong

August 31, 2021

Abstract. Costly manipulation to gain selective admission exhibits strategic complementarity when the admissions quota is loose but strategic substitution when the quota is tight. In a system with two layers of selection, gaming at the university entrance stage can induce a university to give preferential treatment to students from a selective high school, justifying why this school attracts better talent and causing gaming to unravel to the high school entrance stage. We apply this framework to evaluate the impacts of raising the university quota, abolishing university entrance examination, eliminating sorting-by-ability in the high school system, and committing to low-powered selection policies.

Keywords: sequential manipulation; tutoring; ability sorting; equal credibility; preferential treatment; low-powered selection

JEL Classification. D45; D82; J71

*We thank Jimmy Chan, Piotr Dworczak, Navin Kartik, Jin Li, Bruno Strulovici, and participants at the UNSW Organizational Economics Workshop and the Royal Economic Society conference for helpful comments and suggestions. This research is partly funded by the Research Grants Council of Hong Kong (Project No. 17500220).

1. Introduction

Standardized tests results are widely adopted as admissions criteria in education systems around the world. The reliance on test outcomes in the allocation of education resources is often motivated by the desire to select the most talented students to receive these resources. But high-stakes testing also provides powerful inducement for students and their families to engage in costly activities to boost test scores. Private tutoring for test preparation, for example, is prevalent in many countries, especially in Asia.[1] In this paper, we use the terms *tutoring*, *gaming*, or *manipulating* interchangeably as a shorthand for all kinds of costly activities targeted at gaining an advantage in a selective admissions system. In the United States, students can fake disability status to allow them extra time or other accommodations when taking tests.[2] Parents hire coaches to help with packaging their children's college applications. The recent college admissions scandal (Korn and Levitz 2020) testifies to the extreme measures that some parents took to game the system. The goal of this paper is to develop a theoretical framework to study the incentive to engage in these gaming activities in a multi-stage setting, and their implications for selection outcomes.

Private tutoring is so common in Asia that education researchers dub it the "shadow education" system (Bray 1999; Lee, Park and Lee 2009). Although private tutoring may serve useful purposes such as helping the slower students catch up with the demands of formal schooling, the primary motivation behind tutoring seems to be to gain an advantage in competitive examinations. Bray and Kwok (2003) find that the proportion of secondary students in Hong Kong receiving tutoring increases as they reach higher grades. For those in Forms 6 and 7 (equivalent to senior high school), 85% of those receiving tutoring gave "examination preparation" as the main reason. In addition to students' time input, the amount of financial resources that parents spend is huge. A *Financial Times* article estimates that Koreans spend \$20 billion on private tutoring annually.[3] Kaplan, a leading

[1] Park et al. (2016) present data from Program for International Student Assessment surveys to show that the use of tutoring on any subject has increased substantially across a wide range of countries. Outside Asia, heavy users of tutoring (more than 40% of 15-year-olds in 2012) include students from Greece, Hungary, Poland, Russia, Spain, and Turkey. In the U.S., the use of tutoring is not particularly high compared to other countries, but the trend is increasing (from 12% in 2003 to 17% in 2012).

[2] A federal disability designation known as 504 Plan allows students extra time for classroom assignments and standardized tests. Obtaining such a designation requires expensive psychological evaluation. According to a *New York Times* analysis ("Need Extra Time on Tests? It Helps to Have Cash," July 30, 2019), the higher the income, the higher the share of high school students with this designation. In some rich communities, more than one in ten students have such a designation.

[3] "South Korea's Millionaire Tutors," *Financial Times*, 16 June, 2014.

supplier of test preparation services in the U.S., has an annual revenue of \$1.5 billion in 2018. In reaction to the rising trend of test preparation, the College Board redesigned the SAT in 2014 to "rein in the intense coaching and tutoring on how to take the test that often gave affluent students an advantage."[4]

Here we treat gaming activities such as tutoring or buying disability designation as unproductive activities which work by masquerading students as true talent. Critics have long complained that drilling to prepare for standardized examinations and "teaching to the test" may be detrimental to education because they encourage rote learning, and important areas of knowledge are ignored if they are not on the test (e.g., Popham 2001; Volante 2004; see also Lazear 2006 for a different view). The *Financial Times* story mentioned earlier describes an English lesson in a Korean cram school, in which "little English is spoken in the lesson, which comprises an explanation of the TOEIC reading comprehension paper." Some tutoring companies in China even systematically leaked test questions to its students.[5]

Because these gaming activities work through distorting the information available to an admissions system, a main theme of our analysis is how informational externalities affect equilibrium level of gaming and equilibrium admissions standard. We begin with a one-stage model in Section 2. Students can be either of high ability or low ability and they also differ on their costs of gaming the test. A university has limited capacity and prefers to select high-ability students. It does not directly observe students' abilities, but only their test scores, which can be improved by costly gaming. This section shows that tutoring can be strategic complements or strategic substitutes among students, but strategic substitution prevails when competition for university places is intense. This follows because the capacity limit causes diminishing returns to set in as many students get artificially high test scores through gaming. One implication is that a policy that aims to lessen the competition by increasing the university capacity can increase the intensity of gaming.

We build on this model to consider in Section 3 an admissions system with two stages of selection—first at the high school level and then at the university level. Sorting-by-ability is a common feature in many education systems. In the U.S., for example, "magnet" high schools may select students based on entrance examinations. Under such a system, the university can rely on an additional signal, namely whether a student has got into a selective or non-selective high school, to guide its admissions decisions. The objective of

[4] "A New SAT Aims to Realign with Schoolwork," *The New York Times*, March 5, 2014.
[5] "Chinese Education Giant Helps Its Students Game the SAT," Reuters, December 23, 2016.

tutoring in our model is to pool with high-ability students, and the benefit from pooling with these students is larger in the school with more high-ability students. We show that in equilibrium, (1) one high school has a greater fraction of high-ability students than another; (2) the university gives preferential treatment to the better school by admitting its students with a strictly higher probability than students with the same university test score from the other school, even though these two groups of students are equally likely to be of high ability due to more intense gaming in the better school; and (3) the preferential treatment creates a rent from entering the better school, justifying why it can attract better students and causing tutoring to unravel to the high school admissions stage. Interestingly, because the more intense gaming in the better school during the second-stage neutralizes the better initial pool in that school, the final selection outcome for the university only depends on the amount of gaming in the first stage.

Section 4 applies our framework to study a number of issues in selective admissions. Among the more interesting findings are:

1. Abolishing university entrance examination would intensify gaming activity targeting the earlier tests (such as high school entrance examination or high school course grades), and as a result the average quality of the university's admitted students would be lower.
2. When the cost structures of gaming the high school test and gaming the university test are the same, the university admits better students with two-stage selection than with one-stage selection. Therefore abolishing ability-sorting in the high school system worsens the university's selection outcome. However, substantially more resources can be wasted on tutoring under ability-sorting.
3. The university may select better students by committing to a lower-powered admissions policy that is suboptimal ex post but that will mitigate the extent of gaming. But the level of gaming under the optimal policy is positive in both stages, and preferential treatment in favor of students from the selective school remains under the optimal policy.

Our analysis of tutoring and selective admissions is based on a model with very simple test technologies. In Section 5, we show that the main conclusions of our analysis are robust under various generalizations. We can allow finitely many levels of student ability (instead of two levels) and a continuum of test scores (instead of binary scores). The test technology can be stochastic rather than deterministic. In order to illustrate the logic of unraveling in the starkest way, the basic two-stage model has two high schools which are pure labels. A more general model can allow high schools to improve student abilities, with the selective school more able to do so than the non-selective schools. We show that our basic results survive these generalizations. In Section 6, we conclude with some observations concerning possible applications outside the context of college admissions.

Related Literature. Students' manipulation of test scores is a way of lying about their true ability, so our paper is related to the literature on costly lying (Kartik 2009; Kartik, Ottaviani and Squintani 2006). Chan, Li and Suen (2007) consider a signaling model of grade inflation, in which of the cost of signaling is endogenously derived. Frankel and Kartik (2019) study signaling games in which agents have two-dimensional types. When applied to the tutoring example, their model would show how much information is revealed about students' academic ability versus their ability to game the tests. Ball (2021) and Frankel and Kartik (2021) adopt this framework to study design problems, and show how a decision maker can commit to underutilizing data to mitigate gaming and thereby improve allocation efficiency. Unlike Frankel and Kartik (2021), the allocation rule in our model is subject to a capacity constraint. There is also a computer science literature termed "strategic classification" that studies a related problem (Hardt et al. 2016). Different from the existing literature, we consider a two-stage setup, where gaming behavior depends on past outcomes, and the expected gains from gaming unravels to affect the incentive to gain admission to a selective high school. Our focus is on an information receiver that has no commitment ability. Moreover, the capacity constraint of the information receiver creates externality among the information senders.

Our paper shows that the university discriminates based on high school affiliation when there is ability-sorting in the high school system. This aspect is related to the large literature on statistical discrimination. Lundberg and Startz (1983) and Coate and Loury (1993) introduce models in which the labor market incorporates prior beliefs about the abilities of different groups of workers with their test scores to form estimates about their job qualification. In our model, group identity is not fixed, but is the result of first-stage of selection. We show that high schools with a larger pool of high-ability students may be preferentially treated by a college admissions system, despite the fact that in equilibrium their students have the same expected ability as similar students at other schools.

Because competition for limited spots in a university often takes the form of a contest, our research falls into the broad area of rent-seeking contests (Nitzan 1994; Tullock 2001). Siegel (2009) analyzes this type of interactions as an all-pay auction. Unique to our model is that the university does not directly care about how much investments students make in the contest; instead it is concerned about identifying true talents given that some low-ability students are investing to manipulate their test scores. Fang and Noe (2019) use an all-pay auction approach to study selection outcomes. Their focus is on risk-taking behavior in such contests. There is a small theoretical literature on cheating in rank-order tournaments (Curry and Mongrain 2009; Gilpatric 2010; Gilpatric and Reiser 2017). This literature studies how to deter cheating through the design of tournaments, audits, monitoring, and punishments in a single stage setup. Our paper is focused on the information externalities in a multiple-stage contest that aims to learn about the private information of the participants.

2. One Stage of Selection

A large number of students compete for a limited number of university places. For convenience, we model students as a continuum of agents and there are a unit mass of them. The total number (mass) of university places available is $Q \in (0, 1)$, which we refer to as the university quota.

A fraction $\lambda \in (0, 1)$ of the students have high ability. The remaining fraction have low ability. A student knows her true ability but others do not. Students have to sit an entrance examination to gain admission into university. All high-ability students get a high score H. Low-ability students get a low score L if they do not pay for private tutoring. But if a low-ability student pays for costly tutoring, she gets score H in the examination (without changing her true ability). Whether a student has paid for tutoring or not is not observable by the university. We present this simple model with binary types and binary deterministic examination scores in order to illustrate the logic of the analysis in the most transparent way. Section 5 extends the model in various directions, including allowing finitely many types and a continuum of stochastic scores.

A student gets benefit $B > 0$ from being admitted into the university. Since tutoring only affects a low-ability student's score, only low-ability students will buy this service. Among these students, the tutoring cost c is distributed according to distribution F. Assume F has a density that is continuous and everywhere positive on $[0, B]$. A student's payoff is the benefit of getting into the university, if any, minus the cost of tutoring, if any. Because the net payoff is strictly decreasing in c, there is a cutoff cost level S such that a low-ability student prefers to choose tutoring if and only if $c \leq S$. The mass of students who get tutoring is $(1-\lambda)F(S)$.

Given the entrance examination scores, the university selects students to fill its available places. The university's strategy is denoted (X,Y), where X is the probability of admission for high-score students, and Y is the probability of admission for low-score students. We assume that the university prefers admitting a low-ability student to leaving its available spots unfilled. Since $\lambda + (1-\lambda)F(S)$ is the total mass of high scorers, the university's feasibility condition is:

$$(\lambda+(1-\lambda)F(S))X+(1-\lambda)(1-F(S))Y=Q.$$

The university's objective is to maximize the average quality of its student intake, subject to filling its quota according to the feasibility constraint.[6] We assume that the university cannot commit to a selection rule, so its admissions policy has to be optimal given the observed entrance examination results. We will re-visit the no commitment assumption in Section 4.3.

An equilibrium of this model requires: (1) Each low-ability student's tutoring choice maximizes her payoff given the strategies of other students and admissions policy (X,Y). (2) The university's admissions policy is optimal given the observed examination scores and given the students' strategies.

We refer to the posterior probability, given the test score, that a student is of high ability as her *credibility*. Because only low-ability students get a low score, the credibility of a low scorer is 0. Let $K=\lambda/(\lambda+(1-\lambda)F(S))$ denote the credibility of a high scorer. It is optimal for the university to adopt a *priority rule*: high scorers have a strictly higher priority of getting admitted (i.e., $X<1$ implies $Y=0$). If $Q>\lambda+(1-\lambda)F(S)$, we say that the quota is *loose*, in which case all high scorers get a place in the university ($X=1$) and the remaining quota is allocated randomly among low scorers, which gives

$$Y=\frac{Q-(\lambda+(1-\lambda)F(S))}{1-(\lambda+(1-\lambda)F(S))}>0.$$

If $Q\leq\lambda+(1-\lambda)F(S)$, we say that the quota is *tight*, in which case low scorers do not get a place ($Y=0$) and the quota is randomly allocated among high scorers only, which gives

$$X=\frac{Q}{\lambda+(1-\lambda)F(S)}\leq 1.$$

[6] This paper only considers a university whose objective is to select the highest-quality students. Selection on the basis of examination results is the dominant feature of many higher education systems in Asia. In other places a university may have other objectives such as diversity.

Consider a low-ability student's incentive. When the quota is loose, her probability of admission increases from Y to 1 by acquiring tutoring. When the quota is tight, tutoring raises her probability of admission from 0 to X. Therefore her benefit from tutoring is:

$$\beta(S) = \begin{cases} B\left(1 - \frac{Q-(\lambda+(1-\lambda)F(S))}{1-(\lambda+(1-\lambda)F(S))}\right) & \text{if } Q > \lambda + (1-\lambda)F(S), \\ B\left(\frac{Q}{\lambda+(1-\lambda)F(S)} - 0\right) & \text{if } Q \leq \lambda + (1-\lambda)F(S). \end{cases}$$

The benefit from tutoring increases in S when the quota is loose and decreases in S when the quota is tight. When S is small, tutoring exhibits strategic complementarity: as more low-ability students get tutoring, it becomes increasingly difficult for low scorers to gain admission into university, which makes tutoring more valuable. However, tutoring exhibits strategic substitution when the quota is tight: as a lot of students already obtain high scores in the examination through tutoring, the chance of admission for high scorers falls, which lowers the benefit from tutoring. Many discussions on tutoring describe it as an "arms race," in which parents send their children to tutoring schools because their neighbors are doing it.[7] Our analysis suggests that the capacity constraint of a selective admissions system limits the extent of strategic complementarity. When the capacity of the university is reached, diminishing returns to tutoring set in as more and more students are getting high scores. Therefore the self-reinforcing tendency of this "arms race" eventually gives way to a self-limiting tendency to produce an equilibrium level of tutoring.

In the ensuing discussion we focus on the parameter case where the quota is tight. This reflects the view that competition for desirable university places is intense in many societies, so that obtaining high scores alone is not sufficient to guarantee admission. An assumption that will guarantee tight quota for any S is $Q \leq \lambda$. In this case, $\beta(S)$ decreases in S. The following result is immediate, and the proof is omitted.

Proposition 1. *Suppose $Q \leq \lambda$. The unique equilibrium tutoring level $S^* \in (0, B)$ satisfies $BQ/(\lambda + (1-\lambda)F(S^*)) = S^*$.*

The assumption $Q \leq \lambda$ used in Proposition 1 is sufficient but not necessary for equilibrium uniqueness. If $Q > \lambda$ instead, $\beta(S)$ is in general increasing then decreasing in S, as shown in Figure 1. For small quota Q, there is a unique S^* that satisfies $\beta(S^*) = S^*$. When

[7] Ramey and Ramey (2009) titled their paper on intensive parenting for college preparation, "The Rug Rat Race." *The Economist* observes that, in Korea and other Asian economies, "Education has become an arms race in which one parent's additional outlay of time and money forces others to follow suit." ("Will Age Weakens the Asian Tiger Economies?" December 5, 2019.)

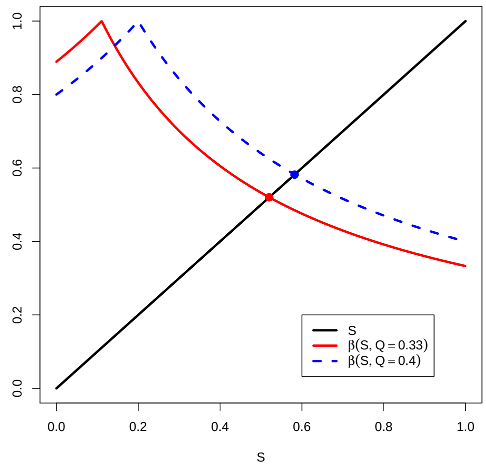

Figure 1. *The benefit from tutoring $\beta(S)$ is increasing then decreasing in S when $Q > \lambda$. An increase in quota Q shifts $\beta(S)$ to the right and increases the equilibrium value of S^* in the largest equilibrium.*

multiple equilibria exist, the quota is tight at the largest equilibrium (i.e., the one with the highest value of S^*), because $\beta(S) > S$ at $Q = \lambda + (1-\lambda)F(S)$ and $\beta(S) < S$ at $S = B$. The largest equilibrium is locally stable in the sense of best-response dynamics. An increase in Q shifts the $\beta(S)$ curve to the right. It is easy to see from the figure that higher Q leads to an increase in S^* in this equilibrium.

Corollary 1. *In the largest equilibrium, the total amount of tutoring S^* increases as the number of university places Q increases.*

Casual reasoning tends to suggest that the availability of more university places would soften the competition for admission and spur less tutoring. This reasoning is incomplete because it ignores the strategic substitution aspect of tutoring under a tight quota. When competition for admission is already quite intense, obtaining a high score in the entrance examination gives one a lottery for entering university but not a guarantee. The availability of more university places increases the chance of admission that comes with this lottery, and therefore raises the benefit of tutoring. The prediction of Corollary 1 is consistent with the experience of Hong Kong. The number of coveted public-funded universities has increased from two before 1991 to nine today, while the number of students competing for these places have fallen due to changes in demographics and the rising trend among students to study abroad. Gaining admission into university changed from a remote possibility into a realistic prospect, particularly if one is willing to spend resources to boost this prospect. Interestingly, the tutoring industry boomed in Hong Kong during the 1990s and remains strong today.[8] Some commentaries in the U.S. tout increasing freshman class sizes at top universities as a way to combat gaming activities.[9] Our analysis suggests that its effect can be the opposite.

3. Two Stages of Selection

In many education systems, selection occurs at multiple transition points of a student's educational progression. For example, competition for admission into New York City's nine specialized high schools is as intense, if not more, as competition for entry into university, and it has spawned a cottage industry of tutoring to prepare for the Specialized High Schools Admissions Test. In some places this type of competition appears to have unraveled to earlier and earlier stages of education. A recent study (Chan et al. 2020) using boundary discontinuity design and cross-boundary matching strategy applied to housing prices suggests that Shanghai parents are willing to pay a 14% premium for an address that would secure the highest priority of getting their children into a local "superstar" primary school. In Hong Kong, a thriving business operates training classes to prepare toddlers for kindergarten interviews.[10]

Building on the model in Section 2, we provide an analysis of the unraveling of tutoring for university admission to tutoring for high school admission in an education system with two stages of selection. The key to this analysis is an informational externality that we will elaborate. To be sure, a selective high school can be different from non-selective ones in terms of the quality of its teachers, the amount of educational resources available, and a myriad of other factors that directly influence students' educational experience and outcomes. We abstract from all these differences in order to emphasize the motive to gain an advantage in college admissions through the informational externality.

Suppose School 1 is a selective high school and has a quota of q_1. School 2 is non-selective and has a quota of $q_2 = 1 - q_1$. All students who do not get into School 1 will attend School 2. For simplicity, we assume that neither high school changes its students' ability. Let λ_1 and λ_2 be the proportion of high-ability students in School 1 and School 2,

[8] Lui and Suen (2005) find that the college premium did not increase over this period, so there is little evidence for an increase in B that might have raised the demand for tutoring.

[9] "Elite-College Admissions are Broken", *The Atlantic*, October 14, 2018.

[10] "Tutoring Centre's Founder Defends Kindergarten Interview Ad Campaign That Went Viral," *South China Morning Post*, 25 May, 2015.

respectively. We have

$$q_1\lambda_1 + (1-q_1)\lambda_2 = \lambda,$$

where λ is the fraction of high-ability students in the population. In this model, which high school a student attends is a pure label that may or may not affect her chances of success at the next stage of selection by university. Section 5 briefly explains why the results will remain valid if the selective high school has an inherent advantage over the non-selective one in improving students' ability.

Our model of competition for admission into School 1 is almost the same as that described in Section 2. Students know their own ability but School 1 does not. There is a high school entrance examination, the result of which is either a high score h or a low score l (we use lowercase letters to denote variables at the first stage of the selection process). High-ability students always get high score h. Low-ability students get score l if they do not pay for tutoring, but will get score h if they pay.

The cost of tutoring (or gaming the admissions system in general) often depends on a student's family background. Richer or more educated parents may incur lower costs due to a lower marginal utility of income or greater knowledge about the various options available in the educational system. We therefore expect that the cost of tutoring is persistent between the two stages of selection. Let the "type" of a low-ability student be represented by c. The cost of tutoring of a type-c student is c at the university entrance stage and is δc at the high school entrance stage. So $\delta < 1$ means that university selection is harder to game than high school selection, and $\delta > 1$ is the opposite case. The distribution of types in the population of low-ability students is given by F on $[0, B]$.

The timing of the two-stage game is as follows. In Stage 1, students decide whether or not to get tutoring. Then high school entrance examination scores realize. School 1 uses these scores to admit students. In Stage 2, students decide whether or not to get tutoring. Then college entrance examination scores realize. The university observes students' high school affiliation and their college entrance examination scores, and uses this information to make admissions decisions. It does not observe students' high school entrance examination scores.

Let s denote the cutoff type who is indifferent between getting tutoring or not in the first stage. School 1 aims to maximize the average ability of its students, subject to the constraint that it fills its entire quota q_1. We denote its admissions policy by (x, y) where x is the probability of admission for students with score h and y is the probability of admission for students with score l. Its feasibility constraint can be written as:

$$(\lambda+(1-\lambda)F(s))x+(1-\lambda)(1-F(s))y=q_1.$$

Students do not directly get any direct benefit from being in a particular high school. They get benefit B from being admitted into university.

For $i=1,2$, denote the cutoff type of low-ability student from School i who pays for tutoring in Stage 2 by C_i. Let the university's admissions policy be represented by (X_1, X_2, Y_1, Y_2), where X_i is the probability that a high scorer from School i is admitted, and Y_i is the probability that a low scorer from School i is admitted. Let $G_i(\cdot)$ denote the distribution of tutoring costs among low-ability students in School i. The objective of the university is to maximize the average quality of the admitted students, subject to the feasibility condition:

$$\sum_{i=1,2} q_i\left[(\lambda_i+(1-\lambda_i)G_i(C_i))X_i+(1-\lambda_i)(1-G_i(C_i))Y_i\right]=Q.$$

An equilibrium of this model requires: (1) Each low-ability student's first and second-stage tutoring choices maximize her payoff given the tutoring choices of other students and given the admissions policies of School 1 and of the university. (2) The admissions policies of School 1 and of the university maximize the quality of their respective student intakes given the information available and given students' tutoring choices.

Assumption 1. $Q<q_1<\lambda$.

We assume that competition for admission into the selective high school and into university is intense. Assumption 1 requires School 1's quota to be tight, implying that students with score h in the high school entrance examination will be rationed (i.e., $x<1$), while students with score l will not have a chance of getting into School 1 ($y=0$). Recall that λ_i denote the proportion of high-ability students from School i ($i=1,2$). By Bayes' rule,

$$\lambda_1=\frac{\lambda}{\lambda+(1-\lambda)F(s)}, \tag{1}$$

and $\lambda_2=(\lambda-q_1\lambda_1)/(1-q_1)$. For any value of s, we have $\lambda_1 \geq \lambda_2$.

If the distribution of students in School 1 and School 2 were exogenous and unrelated to tutoring costs, we would have $G_1=G_2=F$. However, gaming induces disproportionately more students with low tutoring costs to successfully get into the selective School 1. Given that among low-ability students only those with type lower than s have a chance of attending School 1, we obtain

$$G_1(c) = \begin{cases} \frac{F(c)}{F(s)} & \text{for } c \in [0, s), \\ 1 & \text{for } c \in [s, B]. \end{cases}$$

The distribution of types in School 2 is determined by the identity:

$$q_1(1-\lambda_1)G_1(c) + (1-q_1)(1-\lambda_2)G_2(c) = (1-\lambda)F(c), \tag{2}$$

for any $c \in [0, B]$. Clearly, as long as some low-ability students choose not to get tutoring in Stage 1 (i.e., $s < B$), the distribution G_2 first-order stochastically dominates G_1.

Denote the credibility of a high scorer (at the university entrance examination) from School i ($i = 1, 2$) by K_i. The credibility of a low scorer from either school is 0. By Assumption 1, it is not optimal for the university to allocate any of its available places to low scorers; we have $Y_1 = Y_2 = 0$. We say high scorers from School 1 and School 2 are *equally credible* if they are equally likely to be of a high ability, i.e.,

$$\frac{\lambda_1}{\lambda_1 + (1-\lambda_1)G_1(C_1)} = \frac{\lambda_2}{\lambda_2 + (1-\lambda_2)G_2(C_2)}. \tag{3}$$

Lemma 1. *In any second-stage subgame equilibrium with $\lambda_1 > \lambda_2$, (a) high scorers from School 1 and School 2 are equally credible; (b) the fraction of low-ability students who choose tutoring is strictly higher in School 1 than in School 2; and (c) high scorers from School 1 are admitted with a weakly higher probability than those from School 2.*

Proof. (a) Suppose $K_i > K_j$, then high scorers from School i have a higher priority of admission into university than high scorers from School j, i.e., $X_i < 1$ implies $X_j = 0$. If $X_i < 1$, then $X_j = 0$ would imply that no student in School j choose tutoring. Therefore $K_j = 1$, which contradicts $K_i > K_j$. When $X_i = 1$, all students in School i would choose tutoring. If $i = 1$, $X_1 = 1$ would imply that all students in School 1 are admitted into university, which violates $Q < q_1$. If $i = 2$, $X_2 = 1$ would imply $K_2 = \lambda_2 \le \lambda_1 \le K_1$, which contradicts $K_2 > K_1$.

(b) Let $\rho \equiv \lambda_1(1-\lambda_2)/(\lambda_2(1-\lambda_1))$ The equal credibility condition (3) can be written as $\rho = G_1(C_1)/G_2(C_2)$. Therefore, $\rho > 1$ implies $G_1(C_1) > G_2(C_2)$.

(c) Since $C_1 \leq s$, we have $G_1(C_1) = F(C_1)/F(s) = (1-\lambda)\lambda_1 F(C_1)/(\lambda(1-\lambda_1))$. That is,

$$\frac{(1-\lambda)F(C_1)}{G_1(C_1)} - \frac{(1-\lambda_1)\lambda}{\lambda_1} = 0.$$

Therefore,

$$\frac{G_2(C_1)}{G_1(C_1)} - \frac{\lambda_2(1-\lambda_1)}{\lambda_1(1-\lambda_2)} = \frac{1}{(1-q_1)(1-\lambda_2)}\left[\frac{(1-\lambda)F(C_1)}{G_1(C_1)} - \frac{(1-\lambda_1)\lambda}{\lambda_1}\right] = 0.$$

Equal credibility implies:

$$\frac{G_2(C_2)}{G_1(C_1)} - \frac{\lambda_2(1-\lambda_1)}{\lambda_1(1-\lambda_2)} = 0.$$

These two equations imply $G_2(C_1) = G_2(C_2)$, and hence $C_1 = C_2$. If $C_1 = C_2 < s$, the cutoff types must be indifferent between tutoring or not, and therefore $BX_i = C_i$ for $i = 1, 2$ implies $X_1 = X_2$. If $C_1 = C_2 = s$, low-ability students of all types in School 1 prefer tutoring to no tutoring. We have $BX_1 \geq C_1 = C_2 = BX_2$, and therefore $X_1 \geq X_2$. ■

The equal credibility result (Lemma 1(a)) is a consequence of an (ex post) optimal admissions policy, and it holds regardless of whether λ_1 and λ_2 are pre-determined or are the outcome of first-stage selection. One can think of the university as relying on two signals: high school affiliation and university test score. These two signals are sequential in the sense that the students observe the outcome of first (high school affiliation) before deciding on whether to game the second (university test score). The equal credibility result says that, conditional on a positive realization of the second signal, the first signal is not informative to the university. Suppose this is not the case. Then, because of the limited capacity of the university, it will only select students with positive realizations of both signals. This would imply that students have no incentive to game the second signal if the first turns out to be negative, but will still have incentive to game the second signal if the first is positive. But then a negative realization of the first signal actually induces a better belief because the university is assured that the second signal is not gamed. This forms a contradiction.

Whenever there are more high-ability students in the selective school (i.e., $\lambda_1 \geq \lambda_2$), gaming must also be more prevalent in the selective school in order for equal credibility to hold (Lemma 1(b)). If λ_1 and λ_2 are pre-determined and the distributions of types are identical in the two schools, $G_1(C_1) \geq G_2(C_2)$ would immediately imply $C_1 \geq C_2$. When λ_1 and λ_2 are the outcome of first-stage selection, Lemma 1(c) shows that $G_1(C_1) \geq G_2(C_2)$

still implies $C_1 \geq C_2$. Because the cutoff type of student who chooses tutoring is weakly higher in School 1 than in School 2, the indifference condition implies that the university must be admitting high-scorers in School 1 with a weakly higher probability than those in School 2.

Because high scorers from the two schools must be equally credible, any admissions policy (X_1, X_2) is optimal as long as it is feasible. Nevertheless, Lemma 1(c) imposes the restriction $X_1 \geq X_2$ on equilibrium admissions policies. This result holds in any subgame with an arbitrary first-stage cutoff type s. We can strengthen the weak inequality to a strict inequality by considering equilibrium incentives in the first-stage tutoring decision.

Lemma 2. *In any equilibrium, $C_1 = C_2 = s \in (0, B)$ and $X_1 > X_2$.*

Proof. If $s = 0$, then $\lambda_1 = 1$ and $K_1 = 1 > K_2$, which would violate the equal credibility condition. We therefore have $s > 0$. Suppose $C_1 = C_2 < s$. Then a type-s student strictly prefers not to acquire second-stage tutoring no matter she is in School 1 or School 2. Because $Y_1 = Y_2 = 0$, the subgame payoff to type s is 0. This implies that type s strictly prefers not to acquire tutoring in the first stage, a contradiction. This shows that $C_1 = C_2 = s$. Lemma 1 establishes that $X_1 \geq X_2$. If $X_1 = X_2$, however, there is no gain from entering the selective high school, which would violate $s > 0$. We therefore must have $X_1 > X_2$. Finally, if $s = B$, all students in School 1 would have high score in the university entrance examination. This would imply $X_1 \leq Q/q_1 < 1$, and type s would strictly prefer not to pay for tutoring. ■

The highest type of low-ability student in School 1 is type s. Lemma 2 shows that all low-ability students in School 1 choose tutoring in the second stage, but only a fraction of low-ability students in School 2 do so. Despite sorting of high-ability students into the selective School 1, the second-stage gaming decisions neutralize the prior differences between the two schools ($\lambda_1 > \lambda_2$) to make high scorers from both schools equally credible ($K_1 = K_2$). But despite this equal credibility condition, any equilibrium must exhibit preferential treatment in favor of School 1, in that equally credible students are given a strictly higher chance of admission at the favored school ($X_1 > X_2$).

An equilibrium with $X_1 > X_2$ is supported by the fact that $1 = G_1(C_1) > G_2(C_2)$. In this equilibrium, type s in School 1 strictly prefers to get tutoring, while type s in School 2 is indifferent between tutoring or not. With $X_1 > X_2$, there is a rent of $B(X_1 - X_2)$ for both high-ability and low-ability students from being admitted into School 1. This rent provides an incentive for students to compete—by engaging in costly tutoring if necessary—for admission into the selective school, which in turn justifies why $\lambda_1 > \lambda_2$ can arise as an endogenous equilibrium outcome, despite the fact that one's high school affiliation is just a pure label. The incentive to game the university admissions system unravels to the high school admissions stage through this channel.

Proposition 2. *A unique equilibrium exists. Low-ability students with types lower than or equal to s^* pay for tutoring at both stages, where*

$$(\lambda + (1-\lambda)F(s^*))s^* = \frac{BQ}{1+\delta}. \tag{4}$$

No one else pays for tutoring. The equilibrium admissions policy of the university is

$$(X_1^*, X_2^*, Y_1^*, Y_2^*) = \left(\frac{s^*}{B} + \frac{\delta Q}{(1+\delta)q_1}, \frac{s^*}{B}, 0, 0\right).$$

Proof. Lemma 2 implies that type s in School 2 is indifferent between tutoring or not. So we have $X_2 = s/B$. Since $Y_1 = Y_2 = 0$, the admissions probability X_1 is determined by the second-stage feasibility condition:

$$q_1(\lambda_1 + (1-\lambda_1)G_1(C_1))X_1 + (1-q_1)(\lambda_2 + (1-\lambda_2)G_2(C_2))X_2 = Q.$$

Using the equal credibility condition (3) and the fact that $G_1(C_1) = 1$, this reduces to

$$X_1 = \frac{1}{q_1}\left(Q - (\lambda + (1-\lambda)F(s) - q_1)\frac{s}{B}\right).$$

In the first stage, type s can get into School 1 with probability $q_1/(\lambda + (1-\lambda)F(s))$ by investing in first-stage tutoring at cost δs. The indifference condition requires:

$$\delta s = B(X_1 - X_2)\frac{q_1}{\lambda + (1-\lambda)F(s)}.$$

Using the values of X_1 and X_2 derived earlier, we obtain (4), and there is a unique $s^* \in (0, B)$ that satisfies this condition. Plugging $s = s^*$ into X_1 and X_2 gives the equilibrium admissions policy X_1^* and X_2^*. ■

The equilibrium condition (4) in Proposition 2 allows simple comparative statics analysis. As in Corollary 1, a larger university quota Q leads to more gaming at both the university entrance stage and the high school entrance stage. A larger benefit from university admittance has the same effects. Similarly, suppose the cost of gaming falls in the sense of a first-order stochastic decrease in the distribution F. Equation (4) implies that the cutoff s^* decreases, but the fraction of students choosing tutoring $F(s^*)$ increases. As a result, the average quality of students admitted into university will fall.

Proposition 2 establishes that any equilibrium must exhibit preferential treatment in favor of the selective high school, despite the fact that high scorers in both high schools are equally credible. This is because without preferential treatment, the more selective school will have a higher credibility. Preferential treatment induces more gaming in the selective high school to result in equal credibility between the schools.

The logic driving the preferential treatment outcome is different from the discrimination literature, which relies on multiple equilibria. In Coate and Loury (1993), for example, employers set different hiring standards for different groups, which produce different human capital investment incentives for these groups that in turn justify the different standards. In our model, low-ability students in the favored School 1 indeed invest more in tutoring (i.e., $G_1(C_1) > G_2(C_2)$), but such investments reduce rather than enhance credibility for the group as a whole. The preferential-treatment outcome is derived from costly gaming under the equal credibility requirement. Any equilibrium described in Proposition 2 necessarily exhibits discrimination in favor of students from School 1, whereas multiple equilibria in Coate and Loury (1993) is consistent with no discrimination if employers adopt the same standard for both groups. Furthermore, which high school a student belongs to is not exogenously given, but is the outcome of a selection process partly driven by competition for preferential treatment in the university entrance stage.

Our result that that there is a unique equilibrium exhibiting preferential treatment relies on the assumption that School 1 is selective while School 2 is not. Modifying this assumption can give rise to multiple equilibria, but the basic economic forces leading to preferential treatment does not change. Specifically, consider a model where both schools hold entrance examinations and try to recruit students with score h. There are three equilibria: (1) The first equilibrium is identical to the one described in Proposition 2, with the university giving preferential treatment to students from School 1, and with tutoring unraveling to the first stage when students compete to get into this school. Despite the fact that School 2 also tries to select h-scorers, these students will choose to enroll in School 1 if offered admission there. (2) The second equilibrium is the flip side of the first, with

School 2 being the high school which is sought after by students and which is accorded preferential treatment at the second stage. (3) The third equilibrium is one in which students randomly choose between the two schools if offered admission by both schools. The resulting distribution of student ability will be identical, and the equilibrium reduces to the one-stage model described in Proposition 1, with no incentive to get tutoring at the high school entrance stage.

Note, however, that the third (symmetric) equilibrium is not robust in the following sense. Suppose the two high schools are not exactly identical. Say, School 1 has a more beautiful campus that brings an additional payoff $\epsilon > 0$ to its students. No matter how small ϵ is, students who are offered admission by both schools will choose to enroll in School 1, destroying the symmetry in ability distribution between the two schools. On the other hand, the first equilibrium will survive if School 1 has an initial small advantage. To put it slightly differently, small differences in the ex ante quality of the two high schools can lead to intense competition for entry into the selective school that endogenously receives preferential treatment in equilibrium. More interestingly, the second equilibrium, in which School 2 is preferentially treated, will also survive even though it has a small initial disadvantage. This is because the informational externality arising from having a larger fraction of high-ability students is self-reinforcing: if h-scorers choose to join School 2 despite the uglier campus, preferential treatment by the university will produce a rent from joining School 2 that can overcome the ϵ payoff difference to make School 2 more attractive. Therefore, when both high schools can select students, endogenous selection will typically lead to asymmetric distribution of student ability. But which one ends up becoming the "elite school" may be affected more by history-dependent factors than by the underlying quality of the two schools.

4. Applications

4.1. Abolish standardized test requirement for university admission

A growing list of universities in the U.S., including Chicago, Harvard, and Princeton, have recently adopted test-optional admissions policies. The University of California system went further and decided in 2021 to stop considering SAT and ACT scores in its admissions decisions. The arguments for or against standardized tests are multi-faceted. We do not attempt to address all the relevant issues here, but will use the two-stage model of Section 3 to illustrate how abolishing university entrance examination may influence attempts at gaming admissions systems and their effects on selection outcomes. Because there is only one university in our model, considerations arising from competition among universities for high quality applicants are absent from this model.

We use a tilde to denote equilibrium quantities when there is no university entrance examination. Because of the lack of information from standardized test scores H or L, the university's admissions policy is entirely determined by the high school affiliation of the students. When $\tilde{\lambda}_1 > \tilde{\lambda}_2$, the university gives students from School 1 strict priority over those from School 2, so the probability of admission to the university is Q/q_1 for a student from School 1 and is 0 for one from School 2. This implies that the benefit of first-stage tutoring for a low-ability student is:

$$\tilde{\beta}_1(\tilde{s}) = B\left(\frac{Q}{q_1} - 0\right)\frac{q_1}{\lambda + (1-\lambda)F(\tilde{s})}.$$

The following result follows directly from the indifference condition $\delta\tilde{s} = \tilde{\beta}_1(\tilde{s})$, and its proof is omitted.

Proposition 3. *Suppose the university entrance examination is abolished. In the unique equilibrium, all low-ability students with types less than or equal to $\tilde{s}$ acquire tutoring, where*

$$(\lambda + (1-\lambda)F(\tilde{s}))\tilde{s} = \frac{BQ}{\delta}.$$

When there is no university examination, $\tilde{X}_1 = Q/q_1 > X_1^*$ and $\tilde{X}_2 = 0 < X_2^*$. Because the rent from entering School 1 increases from $B(X_1^* - X_2^*)$ to $B(\tilde{X}_1 - \tilde{X}_2)$, competition to get into the selective school becomes more intense, and we have $\tilde{s} > s^*$. Without a university examination, the quality of the university's intake is the same as the quality of an average student from School 1; we have $\tilde{\lambda}_1 = \lambda/(\lambda + (1-\lambda)F(\tilde{s}))$. With an entrance examination, equal credibility implies that the quality of the university's intake is equal to K_1. Since $G_1(C_1) = 1$ in the equilibrium of Proposition 2, we have $K_1 = \lambda_1 = \lambda/(\lambda + (1-\lambda)F(s^*))$. Therefore, $\tilde{s} > s^*$ implies that the quality of the university's intake is strictly lower without a university entrance examination.

Corollary 2. *Abolishing the university entrance examination increases the amount of first-stage gaming and reduces the average ability of the university's admitted students.*

Interestingly, the drop in student quality is not directly due to the university losing a tool to measure student quality. The effect is indirectly caused by more gaming in an earlier stage to distort the remaining tool that the university relies more intensively on: a student's high school affiliation.

4.2. No ability sorting in the high school system

Ability sorting via selective "magnet schools" is a common feature in the education systems of many localities. The merits and demerits of having selective high schools in the system are often the subject of impassioned debates.[11] Our simple model is not equipped to address the complex set of issues related to complementarity in human capital formation, peer effects, race, equity, or social mobility. Less remarked upon in these debates are the informational externalities conferred by elite schools, and the effects of these externalities on college admission outcomes. In this subsection, we use the analysis of Sections 2 and 3 to compare selection outcomes and the extent of gaming in education systems with and without ability sorting across high schools.

Take the one-stage selection model as representing a system where high schools are not differentiated in the ability mix of their student bodies. The average ability of a student admitted by the university is $K = \lambda/(\lambda + (1-\lambda)F(S^*))$, where S^* is the equilibrium level of tutoring described by Proposition 1 in the one-stage setup. On the other hand, ability sorting can be represented by the two-stage model. By the equal credibility result, average ability of the university's student intake with two stages of selection is not improved directly by sorting students into two high schools; it depends solely on the equilibrium level of gaming in the first stage: $K_1 = \lambda/(\lambda+(1-\lambda)F(s^*))$, where s^* is described by Proposition 2. Therefore the comparison of the quality of students admitted into the university boils down to a comparison between S^* and s^*.

The conditions that pin down S^* and s^* are:

$$S = \beta(S) = \frac{Q}{\lambda + (1-\lambda)F(S)} B,$$
$$\delta s = \beta_1(s) = \frac{q_1}{\lambda + (1-\lambda)F(s)} B\left[X_1^*(s) - X_2^*(s)\right].$$

By Proposition 2, $X_2^* > 0$, which in turn implies $X_1^* < Q/q_1$ because the university allocate some of its quota to students from School 2. Therefore, we have $\beta(S) > \beta_1(s)$ if $S = s$. In a one-stage setup, gaming activities target the examination score that the university directly

[11] See, for example, the debate surrounding recent proposals to change the admissions criterion for New York City's specialized high schools, in "What's Going on with New York's Elite Public High Schools?" *The Atlantic*, June 14, 2018.

relies on. In a two-stage setup, the first stage of gaming targets at getting into the elite high school, but the university only partially relies on high school affiliation in making its admissions decision. This lower reliance reduces the benefit to game the admissions decision of the elite high school. If $\delta = 1$, the comparison of $\beta(\cdot)$ against $\beta_1(\cdot)$ immediately implies $S^* \geq s^*$. However, this conclusion can be overturned if the cost of tutoring at the two stages are different, reflecting different levels of difficulties in gaming the two stages of selection. For example, the hourly rate for private tutoring at lower grades tends to be cheaper than that at higher grades. The following result summarizes the discussion; its proof is omitted.

Proposition 4. *There exists a critical value $\hat{\delta} < 1$ such that $s^* > S^*$ and the university's admitted students are of lower average ability under ability sorting (two stages of selection) than under no ability sorting (one stage of selection) if and only if $\delta < \hat{\delta}$.*

Proposition 4 states that ability sorting produces worse selection outcomes when the cost of first-stage tutoring is sufficiently lower than the cost of second-stage tutoring. Another relevant yardstick for comparison is the total expenditure spent on tutoring, which is a wasteful activity in this model. Under two-stage ability sorting, total expenditure (from both stages) on tutoring is:

$$(1+\delta)\int_0^{s^*} c\,\mathrm{d}F(c).$$

In contrast, total expenditure on tutoring with no ability sorting is

$$\int_0^{S^*} c\,\mathrm{d}F(c).$$

If δ is slightly above the critical value $\hat{\delta}$, then s^* would be slightly below S^*, so that ability sorting produces slightly better selection outcomes. But total resources spent on wasteful gaming would be substantially larger (almost by a factor $1+\delta$) under ability sorting.

4.3. Mitigate gaming with commitment

Thus far, we have assumed that the university cannot commit to an admissions policy. Its admissions decisions must be optimal ex post, given its available information and given the strategy of students. A student whose credibility is strictly higher than another student is accorded a strictly higher priority of admission. This means that the university will fully utilize the two available pieces of information: (1) students' high school affiliation, and

(2) university entrance examination scores. This full reliance gives students the incentive to manipulate both sources of information.

In this section, we explore the possibility of using an (ex post) suboptimal admissions policy as a way of controlling the extent of students gaming the selection system. The general idea is that if high reliance on the two pieces of information encourages gaming targeted at them, then relying on them less (by adopting a low-powered selection scheme that differentiates students with different expected ability less) may reduce such activities. To the extent that gaming reduces the efficacy of selection, the university may gain from such measures even if it does not care about the resources spent by students in gaming the selection system.[12]

Viewed in this lens, the policy analyses in Sections 4.1 and 4.2 are special cases of commitments. Section 4.1 is about commitment to no reliance on second-stage test scores, while Section 4.2 is about commitment to no reliance on the high score affiliation. We have shown that both forms of commitment lead to worse selection outcomes (if $\delta = 1$). However, an optimal scheme in which the university can fine-tune its reliance on each piece of information in its admissions may produce strictly better outcomes.

Consider the one-stage model where the university commits to admissions policy (X, Y). The university's problem is:

$$\begin{aligned} \max_{X,Y,S} \quad & X \\ \text{subject to} \quad & (\lambda + (1-\lambda)F(S))X + (1-\lambda)(1-F(S))Y = Q, \\ & B(X - Y) = S, \\ & Y \geq 0. \end{aligned}$$

High-ability students have score H, and the objective is to maximize their chance of admission X. The constraint $B(X - Y) = S$ reflects the response of low-ability students to game the admissions system through costly tutoring. Substituting out the variables X and Y reduces the problem to maximizing $(1 - F(S))S$. Note also that $Y \geq 0$ if and only if

[12] Frankel and Kartik (2021) and Ball (2021) show that a decision maker should commit to underutilize data to reduce manipulation. Whitmeyer (2021) shows that an information receiver can benefit from committing ex ante to observe a noisy signal of the message from the sender. In the finance literature, Goldman and Slezak (2006) point out that high-powered stock option incentives may induce manipulation.

$S \leq S^*$ (where S^* is the equilibrium cutoff in the one-stage model of Section 2). Denote

$$\hat{S} = \underset{S}{\operatorname{argmax}}\,(1-F(S))S.$$

If $\hat{S} \geq S^*$, then the optimal commitment policy and the equilibrium no-commitment policy coincide. If $\hat{S} < S^*$, the optimal policy induces $\hat{S}$ by setting $X < X^*$ and $Y > 0$. That is, the optimal policy does not follow a strict priority rule. The premium associated with high score is lower than the equilibrium premium in this case. Also note that $\hat{S} = 0$ is not optimal, because to eliminate all gaming incentives, the university has to place no reliance on test scores, which wastes valuable information on student ability.

Recall that S^* increases with the size of the quota Q (Corollary 1). If we introduce the assumption that the distribution F satisfies the (weak) monotone hazard rate property, then $(1-F(S))S$ is strictly concave whenever it is increasing in S. Given such an assumption, there exists a critical quota $\hat{Q}$ such that $\hat{S} < S^*$ if and only if $Q > \hat{Q}$. This suggests that commitment ability strictly improves the university's selection outcome if its quota is sufficiently large.

The same logic extends to the two-stage model. Let the university commit to a scheme (X_1, X_2, Y_1, Y_2). Proposition 5 below shows that the university's optimal commitment policy induces students with types below $\hat{s}$ to acquire tutoring at both stages, regardless of whether they are admitted to School 1 or School 2, where

$$\hat{s} = \begin{cases} s^* & \text{if } \hat{S} \geq s^*, \\ \hat{S} & \text{if } \hat{S} < s^*, \end{cases}$$

and s^* is the equilibrium cutoff for the first-stage tutoring decision in Section 3. Whenever $\hat{S} < s^*$ (which is more likely to occur when Q is large), the optimal commitment scheme does not follow a strict priority rule: the university commits to admitting low-score students in School 2 with positive probability (i.e., $Y_2 > 0$), despite a posterior belief that such students have low ability for sure.

Proposition 5. *Suppose F satisfies the monotone hazard property. If the university can commit to an admissions policy (X_1, X_2, Y_1, Y_2), its optimal policy induces the tutoring cutoffs $C_2 = s = \hat{s}$ and all low-ability students in School 1 chooses tutoring. The optimal policy is*

$$Y_2 = Q - \frac{1+\delta}{B}(\lambda + (1-\lambda)F(\hat{s}))\hat{s},$$

and,

$$X_1 = Y_2 + \frac{\hat{s}}{B} + \frac{\delta(Q - Y_2)}{(1+\delta)q_1}, \qquad X_2 = Y_2 + \frac{\hat{s}}{B}, \qquad Y_1 \in \left[0, X_1 - \frac{\hat{s}}{B}\right].$$

The optimal policy does not follow the strict priority rule (i.e., $Y_2 > 0$) if and only if $\hat{s} < s^*$, and it coincides with the no-commitment equilibrium policy if and only if $\hat{s} = s^*$.

In the proof of Proposition 5 (in Appendix B) we first show that a policy that induces no tutoring ($s = 0$) is dominated by a policy that induces the no-commitment equilibrium allocation (which is certainly feasible). We then use the monotone hazard property to ensure that $(1 - F(s))s$ is strictly concave when it is increasing, which in turn implies that $(\lambda + (1-\lambda)F(s))s$ is strictly convex. This assumption, together with Jensen's inequality, can be used to establish that whenever there is a feasible policy such that C_1, C_2 and s are not all equal, there exists another feasible policy with $C_1 = C_2 = s$ which will strictly improve the quality of the university intake. The problem then reduces to finding the value of s that maximizes $(1 - F(s))s$.

Under the optimal commitment policy, a low-ability student in School 2 is indifferent between tutoring and no tutoring if her type is $\hat{s} = B(X_2 - Y_2)$. So $\hat{s} < s^*$ will imply that the premium attached to a high test score H (i.e., $X_2 - Y_2$) is lower under optimal commitment than under no commitment. Similarly a low-ability student is indifferent between first-stage tutoring and no tutoring if her type satisfies $\delta\hat{s}(\lambda + (1-\lambda)F(\hat{s})) = B(X_1 - X_2)q_1$. Again, $\hat{s} < s^*$ will imply that the preferential treatment attached to affiliation with the selective high school (i.e., $X_1 - X_2$) is lower under optimal commitment than under no commitment. A lower-powered selection system improves selection outcomes by mitigating the extent of gaming activities. However, lowering $X_2 - Y_2$ and $X_1 - X_2$ entails raising Y_2 for the quota constraint to be satisfied. This explains why $Y_2 > 0$ under the optimal commitment policy whenever $\hat{s} < s^*$, despite the fact that low-scorers are sure to have low-ability—the optimal commitment policy is ex post suboptimal.[13]

Under the optimal commitment policy, a low-ability student's behavior throughout the two stages does not depend on which high school this student goes to and solely depends on this student's tutoring cost type. Therefore, equal credibility holds under the optimal commitment policy as well: high-score students from School 1 and School 2 are equally likely to be of a high ability.

[13] In Chan and Eyster (2003), the use of non-cutoff admissions policy stems from a preference for diversity and is part of the university's optimal response to a ban on affirmative action.

5. Discussion

We develop two key insights in this paper concerning private tutoring or gaming in college admissions. The first point is that gaming can exhibit both strategic complementarity and strategic substitution, with strategic substitution prevailing when the admissions quota is tight. This leads to the prediction that increasing the number of college places can increase the equilibrium level of tutoring. The second point relates to the unraveling of gaming incentives to earlier stages of the educational progression, leading to competition for entry into selective high schools or primary schools even if those schools are not superior in intrinsic quality to other schools. These points are illustrated in the context of a very simple model with binary ability distribution, binary test scores, and a deterministic gaming technology that always produce a high score for low-ability students. In this section, we provide a more general discussion of our two key results and argue that they are robust to suitable extensions of our basic model.

Strategic substitution under tight quota is a manifestation of the simple fact that gaming does not increase the number of college places available, and therefore it is bound to hit diminishing returns when lots of students are doing it. When the number of college places increase, the benefit from tutoring will reach diminishing returns only when more students are choosing to get tutoring. Thus a larger university quota increases the equilibrium extent of tutoring. In Appendix A, we outline a model to illustrate that this conclusion of Corollary 1 survives in more general environments:

- The ability distribution allows multiple types of student abilities, $a_1 < \ldots < a_N$.
- The test technology is stochastic and test scores are continuously distributed. A student with ability a_n obtains test score $T = a_n + U$, where U is a continuous random variable distributed according to Φ with a log-concave density ϕ.
- More than one type of students can choose tutoring. A student with ability a_n gets test score $T = a_{n+1} + U$ by investing in tutoring.
- The tutoring cost distribution can depend on student ability type in an arbitrary way.

We show in Appendix A that equilibrium in this more general model is characterized by a cutoff test score $\hat{T}$ to satisfy the university's quota constraint. The incentive for a student with ability a_n to get tutoring is

$$\beta(a_n, \hat{T}) = B\left(\Phi(\hat{T} - a_n) - \Phi(\hat{T} - a_{n+1})\right). \tag{5}$$

The relevant observation is that $\beta(a_n, \hat{T})$ is increasing then decreasing in $\hat{T}$ under the log-concavity assumption. When the quota is loose and the admissions threshold $\hat{T}$ is low, a greater extent of gaming would raise the admissions standard and increase the benefit from gaming. However, the opposite effect obtains when the quota is tight and $\hat{T}$ is already high. In this case a greater extent of gaming would further raise the admissions standard and reduce the benefit from gaming (i.e., strategic substitution). This reflects the fact that it becomes increasingly difficult to gain an advantage in meeting the threshold as the standard becomes very high.

In addition, we have assumed in this paper that the benefit from being admitted into university is fixed. If more low-ability students gain entry into university through tutoring, the average quality of university graduates falls and the associated college wage premium may also fall. Thus, the benefit B from entering university may decrease with a greater extent of tutoring. This is an additional force that leads to strategic substitution.

Using the general model in Appendix A to formally establish our second key result (i.e., unraveling) would be more challenging, because keeping track of the progression of the N different types of students is analytically difficult. Nevertheless we can still illustrate the basic intuition using a special case of the model with $N = 2$.

In the model with two types, only the low-ability type (ability a_1) will choose tutoring. To avoid confusion with the label of high schools, we adopt the notation a_L and a_H to stand for a_1 and a_2, respectively. The proportions of low-ability types and high-ability types are $1 - \lambda_i$ and λ_i, respectively, in School i ($i = 1, 2$). The testing technology and the gaming technology at the university entrance stage are the same as described in the beginning of this section.

The university's admissions policy can be summarized by two test score thresholds, (T_1, T_2), for students from Schools 1 and 2. Under this admissions policy, the feasibility condition of the university is:

$$\sum_{i=1,2} q_i \left[(\lambda_i + (1 - \lambda_i) G_i(C_i))(1 - \Phi(T_i - a_H)) + (1 - \lambda_i)(1 - G_i(C_i))(1 - \Phi(T_i - a_L)) \right] = Q. \tag{6}$$

In the above expression, the distribution of cost types G_i among low-ability students depends on the first-stage selection process. We describe this process more fully in Appendix A. For the purpose of this discussion, we note that the first-stage tutoring decision is characterized by a cutoff cost type s, below which low-ability students will choose tutoring. If only students who choose tutoring could have a chance of getting into School 1, we would have $G_1(c) = F(c)/F(s)$ for $c \in [0, s]$, as in the basic model with binary scores. But when high-school entrance test scores are stochastic, some low-ability students with cost above s will be selected by School 1. Hence, $G_1(\cdot)$ first-order stochastically dominates $F(\cdot)/F(s)$.

In this model with continuous scores, the optimal admissions policy requires the university to equalize the credibility of the *marginal student* from the two schools. Let $K_i(T)$ represent the posterior probability that a student from School i with test score T is a high-ability type. The equal credibility condition require $K_1(T_1) = K_2(T_2)$, which can be expressed (in odds ratio form) as:

$$\frac{\lambda_1}{1-\lambda_1}\frac{\ell(T_1)}{G_1(C_1)\ell(T_1)+1-G_1(C_1)} = \frac{\lambda_2}{1-\lambda_2}\frac{\ell(T_2)}{G_2(C_2)\ell(T_2)+1-G_2(C_2)}, \tag{7}$$

where $\ell(T_i) = \phi(T_i - a_H)/\phi(T_i - a_L)$ is the relative likelihood that a student with test score T_i is a high-ability rather than a low-ability student. If (7) fails, the university could improve the quality of its intake by lowering the cutoff score at the school whose marginal students are more credible, while raising the cutoff score at the other school. In equilibrium, the admissions policy (T_1, T_2) satisfies equations (6) and (7), and the second-stage cutoffs satisfy $C_i = \beta(a_L, T_i)$ for $i = 1, 2$, where $\beta(\cdot)$ is given by (5).

The following result is key to the unraveling of gaming incentives.

Proposition 6. *In any second-stage subgame equilibrium with $\lambda_1 > \lambda_2$, the admissions standard for the selective school is strictly lower, i.e., $T_1 < T_2$.*

To understand this result, suppose the quota is tight so that both T_1 and T_2 are quite high, with $\ell(T_i) \geq 1$. (The proof of Proposition 6 in Appendix A considers the other cases too.) Then the equal credibility condition (7) requires either $\ell(T_1) < \ell(T_2)$ or $G_1(C_1) > G_2(C_2)$. Intuitively, if there were no gaming of test scores, standard Bayesian reasoning suggests that the university should pay attention to both prior knowledge about the quality of the high schools and information from test scores, and therefore $\lambda_1 > \lambda_2$ would imply $T_1 < T_2$. The logic of this result has been pointed out, for example, in the literature on optimal judicial standards (Farmer and Terrell 2001). When students can manipulate their test scores, students from School 1 can in principle become equally credible as those from School 2 if a sufficiently large fraction $G_1(C_1)$ of its low-ability students choose gaming. But imperfect selection at the high-school entrance stage implies that $G_1(C_1)$ cannot be larger than $F(C_1)/F(s)$ due to first-order stochastic dominance. Even when $G_1(C_1)$ reaches this upper bound, we can use the same argument as that provided in Lemma 1 to show that equal credibility cannot be attained with $\ell(T_1) \geq \ell(T_2)$. So equal credibility implies that T_1 must be strictly lower than T_2, meaning that students from the selective School 1 are "treated preferentially" than those from School 2.

Because any subgame exhibits preferential treatment in favor of the selective school, a student will expect to get admitted into university with a higher probability if she is in School 1 than if she is in School 2. For a high-ability student, the rent from entering School 1 is $B(\Phi(T_2 - a_H) - \Phi(T_1 - a_H))$. For a low-ability student of type c, the rent is

$$B(\Phi(T_2 - a_L) - \Phi(T_1 - a_L)) + \max\{\beta(a_L, T_1) - c, 0\}.$$

The prospect of earning this rent induces competition to enter School 1, so that tutoring unravels to the high school entrance stage and $\lambda_1 > \lambda_2$ arises as an equilibrium outcome, in much the same way as that described in the binary score model.

In Appendix A we show that when the university quota is tight, there exists an equilibrium of the two-stage model with $T_2^* > T_1^*$. Strategic substitution implies that that $\beta(a_L, \cdot)$ is decreasing when the quota is tight. So $T_2^* > T_1^*$ implies $C_1^* > C_2^*$. Because School 1 tends to select students with low tutoring costs, G_1 is first-order stochastically dominated by G_2. Therefore, $G_1(C_1^*) > G_1(C_2^*) > G_2(C_2^*)$, meaning that low-ability students in School 1 are more likely to game the admissions system than those in School 2. The application success rate is higher at the selective School 1 for three reasons. First, the admissions standard is lower. Second, a greater proportion of its students have high ability. Third, a greater fraction of its low-ability students choose tutoring to mimic the high-ability students. These conclusions from the continuous-scores model are similar to those obtained from the basic binary-scores model.

Proposition 6 also holds if we let School 1 to be better at bringing some inherent value to the students. For example, suppose School 1 transforms ϵ_1 share of its a_L students into a_H students, while School 2 transforms ϵ_2 share of its a_L students in to a_H students, where $1 > \epsilon_1 \geq \epsilon_2 \geq 0$. The transformation is independent of the cost type of the students, and the students who gain a higher ability is aware of it before deciding whether to game the university test.

With this change, equal credibility still holds, but now the credibility is calculated based on the post-transformation proportion of high-ability students in each school, instead of the initial proportions. Let $\bar{\lambda}_i$ $(i = 1, 2)$ denote the proportion of high-ability students in School i after the transformation has happened. Immediately, we have $\bar{\lambda}_i \geq \lambda_i$, and $\lambda_1 > \lambda_2$ implies $\bar{\lambda}_1 > \bar{\lambda}_2$. Then the proof of Proposition 6 goes through verbatim with λ_i replaced by $\bar{\lambda}_i$.

6. Conclusion

Information manipulation is common in many aspects of life when the information collected has implications for resource allocation. Creative accounting, fake product reviews, and credit score management are among the more important examples. In this paper we focus on information manipulation in college admissions, partly because of its significance to students and their families, and partly because private tutoring for test preparation is such a pervasive feature of the education system, especially in Asia where high-stakes testing plays a prominent role in resource allocation in education. More generally, tutoring or buying disability designation resembles an "influence activity" as described by Milgrom (1988) and Milgrom and Roberts (1988). Much of this literature in organizational economics (e.g., Prendergast and Topel 1996; Ederer, Holden and Meyer 2018; Li, Mukherjee and Vasconcelos 2021) focuses on individual behavior: how individuals respond to incentives in unproductive ways and how to design contracts to reduce gaming. It will be interesting to study how informational externality among individuals and the sequential nature of the selection process affect organizational outcomes, as we do in this paper in the context of a selective admissions system.

Appendix

A. A model with more general environment

We can generalize the one-stage setup of Section 2 to allow finitely many types, a continuum of scores, type-dependent tutoring costs, and uncertainty in what score a student will get in the test. Here, all ability types may engage in tutoring to different extent. We show that the main insight of Section 2—that a higher university quota will increase the amount of tutoring—still holds in this more general setup.

Let student ability be $a \in \{a_n\}_{n=1}^N$ with $a_n < a_{n'}$ for any $n < n'$. Normalize the total mass of students to be 1. The mass of students with ability a_n is λ_n. A higher ability student tends to get a higher score. A student with ability a_n obtains test core $T = a_n + U$ if she does not get tutoring, where U is a continuous random variable with distribution Φ on $(-\infty, \infty)$. We assume that the corresponding density function ϕ is strictly log-concave. This assumption ensures that, for $T' > T$, the likelihood ratio, $\phi(T' - a)/\phi(T - a)$ strictly increases in a.

Tutoring allows a_n to mimic the next higher type a_{n+1} for $n < N$. That is, type a_n gets score $T = a_{n+1} + U$ with tutoring. The tutoring cost c is distributed according to distribution F_n, which has a continuous and everywhere positive density function on $[0, B]$. We allow the cost distribution F_n to depend on type in an arbitrary way.

Let S_n denote the cutoff cost level such that students with cost $c \leq S_n$ and ability $a = a_n$ gets tutoring. The likelihood ratio of obtaining test score T for type a_{n+1} relative to type a_n is:

$$\frac{(1 - F_{n+1}(S_{n+1}))\,\phi(T - a_{n+1}) + F_{n+1}(S_{n+1})\phi(T - a_{n+2})}{(1 - F_n(S_n))\,\phi(T - a_n) + F_n(S_n)\phi(T - a_{n+1})}.$$

This likelihood ratio is increasing in T because both $\phi(T - a_{n+1})/\phi(T - a_n)$ and $\phi(T - a_{n+2})/\phi(T - a_{n+1})$ increase in T. Thus, regardless of the tutoring behaviors, the posterior belief about student ability is strictly increasing (in the sense of likelihood-ratio dominance) in the test score. The optimal admissions policy for a university that wants to maximize the average ability of its intake is a cutoff rule—a student is admitted if and only if her test score exceeds some cutoff $\hat{T}$. For a student with ability a_n, the benefit from tutoring is:

$$\beta(a_n, \hat{T}) = B\left(\Phi(\hat{T} - a_n) - \Phi(\hat{T} - a_{n+1})\right).$$

By log-concavity of ϕ, $\beta(a_n, \cdot)$ is first increasing and then decreasing in $\hat{T}$ for any a_n. For sufficiently large $\hat{T}$, the partial derivative with respect to $\hat{T}$ is negative.

For the cutoff rule $\hat{T}$ to satisfy the university's quota constraint, we require the mass of students with scores above the cutoff to be equal to the quota:

$$\sum_n \left[F_n\left(\beta(a_n,\hat{T})\right)\left(1-\Phi(\hat{T}-a_{n+1})\right)+\left(1-F_n\left(\beta(a_n,\hat{T})\right)\right)\left(1-\Phi(\hat{T}-a_n)\right)\right]\lambda_n=Q,$$

where we use the equilibrium restriction that $S_n = \beta(a_n,\hat{T})$. The left-hand-side of the above equation goes to 1 as $\hat{T}\to-\infty$ and goes to 0 as $\hat{T}\to\infty$. Therefore, there exists $\hat{T}$ that satisfies the equation. Consider the largest such solution among all solutions and denote it by $\hat{T}^*$. At such $\hat{T}^*$, the left-hand-side of the quota constraint crosses Q from above, and so we have $\partial\hat{T}^*/\partial Q<0$. Moreover, $\hat{T}^*$ approaches infinity as Q approaches 0.

The equilibrium measure of students getting tutoring is:

$$\sum_n F_n\left(\beta(a_n,\hat{T}^*)\right)\lambda_n.$$

Because $\hat{T}^*$ is decreasing in Q and $\beta(a_n,\cdot)$ is decreasing in $\hat{T}^*$ for Q sufficiently small, the measure of students getting tutoring is increasing in Q when competition for college admission is intense. Therefore, we have the counterpart to Corollary 1: in the largest equilibrium, the total amount of tutoring increases as the number of university places increases.

For the analysis of two stages of gaming, we restrict the general model to the case of $N=2$ and let $a_L=a_1$, $a_H=a_2$. The proportion of high-ability students in School i ($i=1,2$) is denoted λ_i. We use $\ell(T)=\phi(T-a_H)/\phi(T-a_L)$ to denote the likelihood ratio. We use T^0 to denote the test score that satisfies $\ell(T^0)=1$.

For the first stage of selection, let test scores in the high-school entrance examination be given by $t=a_j+u$ for $j=L,H$, with u being a random variable distributed according to Ψ, with a log-concave density ψ. If a low-ability students of cost type c pays to get first-stage tutoring at cost δc, her test score will be $t=a_H+u$. The cost distribution among all low-ability students is F on $[0,B]$. We continue to use s to represent the first-stage cutoff type. If the admission threshold for School 1 in the high-school entrance examination is $\hat{t}$, we have

$$G_1(c)=\begin{cases}\frac{F(c)(1-\Psi(\hat{t}-a_H))}{F(s)(1-\Psi(\hat{t}-a_H))+(1-F(s))(1-\Psi(\hat{t}-a_L))} & \text{if } c\in[0,s),\\ \frac{F(s)(1-\Psi(\hat{t}-a_H))+(F(c)-F(s))(1-\Psi(\hat{t}-a_L))}{F(s)(1-\Psi(\hat{t}-a_H))+(1-F(s))(1-\Psi(\hat{t}-a_L))} & \text{if } c\in[s,B].\end{cases}$$

Observe that $G_1(\cdot)$ first-order stochastically dominates $\max\{F(\cdot)/F(s), 1\}$.

Proposition 6 of the text establishes that $\lambda_1 > \lambda_2$ implies $T_1 < T_2$, where T_i is the university admissions threshold for students from School i. We provide a proof of the proposition here.

Proof of Proposition 6. Suppose to the contrary that $T_1 \geq T_2$. There are three cases.

(a) If $T_1 \geq T^0 \geq T_2$, then $K_1(T_1) \geq \lambda_1 > \lambda_2 \geq K_2(T_2)$, violating the equal credibility condition.

(b) If $T^0 \geq T_1 \geq T_2$, then $1 \geq \ell(T_1) \geq \ell(T_2)$. Because $\beta(a_L, \cdot)$ is increasing for such values of T_i, we have $\beta(a_L, T_1) \geq \beta(a_L, T_2)$. The indifference conditions at the two schools then imply $C_1 \geq C_2$. Furthermore, since School 1 tends to select students with low tutoring costs, G_2 first-order stochastically dominates G_1, and therefore $G_1(C_1) \geq G_1(C_2) \geq G_2(C_2)$. Together, these inequalities imply

$$\frac{\ell(T_1)}{G_1(C_1)\ell(T_1) + 1 - G_1(C_1)} \geq \frac{\ell(T_2)}{G_2(C_2)\ell(T_2) + 1 - G_2(C_2)},$$

which contradicts (7) whenever $\lambda_1 > \lambda_2$.

(c) If $T_1 \geq T_2 \geq T^0$, then we have $\ell(T_1) > \ell(T_2)$. Recall that we define $\rho = \lambda_1(1 - \lambda_2)/(\lambda_2(1 - \lambda_1))$. The equal credibility condition (7) implies

$$\rho = \frac{\ell(T_2)}{G_2(C_2)\ell(T_2) + 1 - G_2(C_2)} \frac{G_1(C_1)\ell(T_1) + 1 - G_1(C_1)}{\ell(T_1)} < \frac{G_1(C_1)\ell(T_1) + 1 - G_1(C_1)}{G_2(C_2)\ell(T_1) + 1 - G_2(C_2)}.$$

Since $\ell(T_1) \geq 1$ and $\rho > 1$, the above inequality implies $G_1(C_1) > G_2(C_2)$. Therefore,

$$\rho < \frac{G_1(C_1)\ell(T_1) + 1 - G_1(C_1)}{G_2(C_2)\ell(T_1) + 1 - G_2(C_2)} < \frac{G_1(C_1)}{G_2(C_2)}.$$

For $T_i \geq T^0$, $\beta(a_L, \cdot)$ is decreasing in the relevant region. So $T_1 \geq T_2$ implies $\beta(a_L, T_1) \leq \beta(a_L, T_2)$, which in turn implies $C_1 \leq C_2$. We therefore obtain $\rho < G_1(C_1)/G_2(C_1)$. However, we have $G_1(C_1) \leq F(C_1)/F(s)$ due to first-order stochastic dominance. By the proof of Lemma 1(c), this is equivalent to $\rho \geq G_1(C_1)/G_2(C_1)$, which leads to a contradiction. ■

Proposition 6 is premised on the condition that $\lambda_1 > \lambda_2$. Despite the fact that the two high schools have identical intrinsic quality, the following proposition shows that there exists an equilibrium in which endogenous differences between the two high schools can arise, so that $\lambda_1 > \lambda_2$ indeed arise as an equilibrium outcome, as gaming for selective admission unravels to the high-school entrance stage. The condition stated in Proposition 7 below means that if the university recruits exclusively from School 1, its standard must be higher than T^0 even if no student in that school chooses tutoring. This condition implies that any feasible admissions standard must satisfy $T_1 > T^0$, and by Proposition 6 we must have $T_2 \geq T_1 > T^0$ in equilibrium. It serves a similar purpose as Assumption 1 does for the binary score model, namely to ensure that strategic substitution prevails in the relevant region.

Proposition 7. *If $Q < q_1\left[\lambda(1-\Phi(T^0-a_H))+(1-\lambda)(1-\Phi(T^0-a_L))\right]$, there exists an equilibrium with $T_2^* > T_1^* > T^0$. In this equilibrium, gaming is more prevalent in the selective school, and the university application success rate is also higher in the selective school.*

Proof. Fix a first-stage cutoff $s \in [0, B]$. For such s, there is a unique admissions standard t_0 that will fill the quota for School 1, given by:

$$(\lambda+(1-\lambda)F(s))(1-\Psi(t_0-a_H))+(1-\lambda)(1-F(s))(1-\Psi(t_0-a_L)) = q_1.$$

This gives $\lambda_1 = \lambda(1-\Psi(t_0-a_H))/q_1$ and $\lambda_2 = (\lambda - q_1\lambda_1)/(1-q_1) < \lambda_1$. For such λ_1, λ_2, and s, the second-stage subgame must satisfy (6), (7), and the indifference conditions $C_i = \beta(a_L, T_i)$ for $i = 1, 2$.

Consider a university policy $(T_1, T_2) \in [T^0, 1]^2$. The indifference conditions determine C_i as a decreasing function of T_i. As T_1 increases, C_1 falls, and the university will admit fewer students from School 1. To satisfy the quota constraint (6), it must lower T_2 (which will induce a corresponding increase in C_2). Equation (6) therefore defines an implicit function, $T_2 = \tau_{\text{fea}}(T_1)$, which is downward-sloping. At T_1, if the university quota would be exceeded even when $T_2 = \infty$, we define $\tau_{\text{fea}}(T_1)$ to be equal to ∞. The assumption stated in the proposition implies $\tau_{\text{fea}}(T^0) = \infty$. Obviously if T_1 goes to infinity, T_1 must be finite because no student would be admitted to university otherwise. We thus have $\lim_{T_1\to\infty} \tau_{\text{fea}}(T_1) < \infty$.

As T_1 increases and C_1 falls, the marginally admitted student from School 1 becomes more credible. To maintain equal credibility across the two schools requires T_2 to rise (which will induce C_2 to fall). Equation (7) therefore defines an implicit function, $T_2 = \tau_{\text{cre}}(T_1)$, which is upward-sloping, with $\tau_{\text{cre}}(T^0) < \infty$ and $\lim_{T_1\to\infty} \tau_{\text{cre}}(T_1) = \infty$. This establishes that, for each s, there exists a unique pair (T_1, T_2) such that $T_2 = \tau_{\text{cre}}(T_1)$ and

$T_2 = \tau_{\text{fea}}(T_1)$. By Proposition 6, $T_2 > T_1 > T^0$.

For each s, the corresponding standard $\hat{t}$ for selection into School 1 is determined by:

$$(\lambda + (1-\lambda)F(s))(1-\Psi(\hat{t}-a_H)) + (1-\lambda)(1-F(s))(1-\Psi(\hat{t}-a_L)) = q_1.$$

Define $\sigma(s)$ to be equal to:

$$\min\left\{\frac{1}{\delta}(\Psi(\hat{t}-a_L)-\Psi(\hat{t}-a_H))\left[B(\Phi(T_2-a_L)-\Phi(T_1-a_L))+\max\{\beta_L(a_L,T_1)-s,0\}\right],B\right\},$$

where $\hat{t}$, T_1 and T_2 are determined by s. By investing in first-stage tutoring, a low-ability raises her chance of being admitted into School 1 from $1-\Psi(\hat{t}-a_L)$ to $1-\Psi(\hat{t}-a_H)$. The rent from being admitted to School 1 is given by the term is square brackets and is positive. Therefore, $\sigma(\cdot)$ is a continuous mapping from $[0,B]$ to itself. There exists $s^* \in [0,B]$ such that $s^* = \sigma(s^*)$. The admissions standards $\hat{t}^*$, T_1^* and T_2^* corresponding to such cutoff s^*, together with the associated second-stage cutoffs C_1^* and C_2^* determined by the indifference conditions, constitute an equilibrium of the model. ■

B. Proof of Proposition 5

The university chooses $(X_1, X_2, Y_1, Y_2) \in [0,1]^4$ to maximize the mass of high ability students admitted: $V = q_1\lambda_1 X_1 + (1-q_1)\lambda_2 X_2$ subject to the feasibility constraint. Because $F(\cdot)$ satisfies monotone hazard rate, $(1-F(s))s$ is strictly concave over $(0,\hat{S})$, where $\hat{S}$ is the (unique) maximizer of $(1-F(s))s$. It follows that $\hat{S} \in (0,B)$. In the following, we let $\theta(s) \equiv (1-F(s))s$ and $\Theta(s) \equiv (\lambda + (1-\lambda)F(s))s$. Therefore, $\theta(\cdot)$ is strictly concave on $(0,\hat{S})$ and linear on $(-\infty,0)$, while $\Theta(\cdot)$ is strictly convex on $(0,\hat{S})$ and linear on $(-\infty,0)$.

Let s denote the cost type that is indifferent between tutoring and not in the first stage. Let C_i denote the cost type that is indifferent between tutoring and not in the second stage in School i $(i=1,2)$. Note that, for a slight abuse of notation, in this proof we allow s, C_2 to be outside the range $[0,B]$ and allow C_1 to be outside the range $[0,s]$ because they are *indifferent types* rather than the *cutoff types*. For example, if the indifferent type $C_1 < 0$, then $F(C_1) = 0$. As will be shown later, at the optimal policy the indifferent type and the cutoff type may only differ when the indifferent type is above the cutoff type in School 1.

We first show that a policy that results in no first-stage tutoring is not optimal.

Claim. Under the optimal policy, $s > 0$.

Proof of claim. We establish this claim by showing that $X_1 > X_2$ under the optimal policy, so the optimal policy will induce an indifferent type with positive cost to choose tutoring in the first stage.

Step 1. We show that among policies with $X_1 \leq X_2$, the policy $X_1 = X_2$ is optimal. When $X_1 \leq X_2$, all low-ability students go to School 2, so we have $G_2(C_2) = F(C_2)$ and $\lambda_1 = 1$, the best policy with $X_1 \leq X_2$ solves:

$$\begin{aligned}
\max_{X_1, X_2, C_2} \quad & q_1 X_1 + (\lambda - q_1) X_2 \\
\text{subject to} \quad & q_1 X_1 + (1-q_1) X_2 - (1-q_1)(1-\lambda_2)(1-F(C_2))\frac{C_2}{B} = Q, \\
& X_2 \geq \frac{C_2}{B}, \\
& X_1 \leq X_2.
\end{aligned}$$

Suppose $X_1 < X_2$ at the optimal solution, then it must be that $X_2 = C_2/B$, because otherwise, one can reduce X_2 by $q_1\epsilon$ and increase X_1 by $(1-q_1)\epsilon$ with positive ϵ small enough to improve the objective function value and keep the constraints all satisfied. Therefore, $X_1 < X_2$ implies $X_2 = C_2/B$. Then, $X_1 < C_2/B$ must be the solution to:

$$\begin{aligned}
\max_{X_1, C_2} \quad & q_1 B X_1 + (\lambda - q_1) C_2 \\
\text{subject to} \quad & q_1 B X_1 + (1-q_1) C_2 - (1-q_1)(1-\lambda_2)(1-F(C_2)) C_2 = BQ, \\
& B X_1 \leq C_2.
\end{aligned}$$

This further simplifies to:

$$\begin{aligned}
\max_{C_2} \quad & BQ - (1-\lambda) F(C_2) C_2 \\
\text{subject to} \quad & (\lambda + (1-\lambda) F(C_2)) C_2 \geq 0.
\end{aligned}$$

Because the objective function is decreasing in C_2, the constraint has to be binding, which contradicts $X_1 < C_2/B$.

Step 2. We next show that a policy with $X_1 = X_2$ is suboptimal. Because $X_1 = X_2 = C_2/B + Y_2$, we have $V = (\lambda/B)(C_2 + BY_2)$, where C_2 and Y_2 are subject to the feasibility constraint:

$$q_1 C_2 + (1-q_1)(\lambda_2 + (1-\lambda_2) G_2(C_2)) C_2 + BY_2 = BQ.$$

Because $G_2(C_2) = F(C_2)$ and $\lambda_1 = 1$, the feasibility constraint simplifies to:

$$C_2 + BY_2 - (1-\lambda)(1-F(C_2))C_2 = BQ.$$

Let V^* represent the value of the objective function when $C_1 = C_2 = s = s^*$. Then,

$$\begin{aligned} B(V^* - V) &= \lambda[(1+\delta)s^* - (C_2 + BY_2)] \\ &= \lambda[(1+\delta)s^* - BQ - (1-\lambda)(1-F(C_2))C_2] \\ &= \lambda[(1+\delta)s^* - (1+\delta)s^*(\lambda + (1-\lambda)F(s^*)) - (1-\lambda)(1-F(C_2))C_2] \\ &= \lambda(1-\lambda)[(1+\delta)(1-F(s^*))s^* - (1-F(C_2))C_2]. \end{aligned}$$

(i) If $C_2 \le s^* \le \hat{S}$, then because $(1-F(s))s$ is strictly increasing over $[0,\hat{S}]$, $(1-F(s^*))s^* \le (1-F(C_2))C_2$, which implies $V^* > V$. (ii) If $s^* < C_2 \le (1+\delta)s^*$, it also implies $V^* > V$. (iii) If $(1+\delta)s^* < C_2$, then $C_2-(1-\lambda)(1-F(C_2))C_2 = (\lambda+(1-\lambda)F(C_2))C_2 > (1+\delta)(\lambda+(1-\lambda)F(s^*))s^* = BQ$, which contradicts the feasibility constraint. So case (iii) is not possible. The remaining case is (iv) $s^* > \hat{S}$, in which case the equilibrium allocation has $C_1 = C_2 = \hat{S}$. This will result in $B(V^* - V) = \lambda(1-\lambda)[(1+\delta)\theta(\hat{S}) - \theta(C_2)] > 0$. Therefore, in all cases the policy $X_1 = X_2$ is strictly dominated by the no-commitment equilibrium allocation, and hence cannot be optimal. □

Let p denote the probability that a low-ability student with tutoring get admitted into School 1, which depends on the policy. If $C_1 < C_2$, the net benefit from first-stage tutoring for type c is

$$\begin{cases} p(BX_1 - BX_2) - \delta c & \text{if } c < C_1, \\ p(BY_1 - (BX_2 - c)) - \delta c & \text{if } c \in [C_1, C_2], \\ p(BY_1 - BY_2) - \delta c & \text{if } c > C_2. \end{cases}$$

The benefit is non-monotone in c if $\delta < p$. In this case, we need to consider "non-cutoff policies," where the set of types who obtains tutoring in Stage 1 may take the form $[0,\underline{s}] \cup [s,\bar{s}]$. If $C_1 \ge C_2$ or $\delta > p$, then all policies are "cutoff policies," where types below s choose tutoring in Stage 1. We consider cutoff policies first.

Cutoff policies. There are four possible types of cutoff policies: (1) $C_1 \le s$ and $C_2 \ge s$; (2) $C_1 \le s$ and $C_2 \le s$; (3) $C_1 \ge s$ and $C_2 \ge s$; and (4) $C_1 \ge s$ and $C_2 \le s$.

Consider case (1), $C_1 \le s$ and $C_2 \ge s$, where we restrict attention to parameters such that $\delta \ge p$ because we are dealing with cutoff policies.

The feasibility condition is:

$$q_1(\lambda_1+(1-\lambda_1)G_1(C_1))X_1+q_1(1-\lambda_1)(1-G_1(C_1))Y_1$$
$$+(1-q_1)(\lambda_2+(1-\lambda_2)G_2(C_2))X_2+(1-q_2)(1-\lambda_2)(1-G_2(C_2))Y_2=Q.$$

Note that $X_2=Y_2+C_2/B$ and $X_1=Y_1+C_1/B$. The feasibility condition becomes:

$$q_1BY_1+q_1(\lambda_1+(1-\lambda_1)G_1(C_1))C_1$$
$$+(1-q_1)BY_2+(1-q_1)(\lambda_2+(1-\lambda_2)G_2(C_2))C_2=BQ.$$

The indifference condition in the first stage for type s is:

$$\delta s=\frac{q_1}{\lambda+(1-\lambda)F(s)}(BY_1-BX_2+s).$$

Plugging in $X_2=Y_2+C_2/B$, the indifference condition gives:

$$q_1(BY_1-BY_2)=\delta(\lambda+(1-\lambda)F(s))s+q_1C_2-q_1s.$$

Incorporating this equation into the feasibility condition gives:

$$BY_2+\delta(\lambda+(1-\lambda)F(s))s+C_2-q_1s+q_1(\lambda_1+(1-\lambda_1)G_1(C_1))C_1-(1-q_1)(1-\lambda_2)(1-G_2(C_2))C_2=BQ.$$

Because $C_1\leq s$ and the first period has a cutoff nature, we have $G_1(C_1)=F(C_1)/F(s)$ and $G_2(C_2)=[F(C_2)-pF(s)]/[1-pF(s)]$, where $p=q_1/(\lambda+(1-\lambda)F(s))$ is the probability of admission into School 1. This implies

$$q_1(\lambda_1+(1-\lambda_1)G_1(C_1))=p(\lambda+(1-\lambda)F(C_1)),$$
$$(1-q_1)(1-\lambda_2)(1-G_2(C_2))=(1-\lambda)(1-F(C_2)).$$

Plugging these two equations into the feasibility condition and upon simplifying, we have:

$$BY_2+(\delta-p)(\lambda+(1-\lambda)F(s))s+p(\lambda+(1-\lambda)F(C_1))C_1+(\lambda+(1-\lambda)F(C_2))C_2=BQ.$$

Therefore, the constraint $Y_2\geq 0$ is equivalent to:

$$BQ-(\delta-p)\Theta(s)-p\Theta(C_1)-\Theta(C_2)\geq 0.$$

Similarly, plugging $X_i = Y_i + C_i/B$ $(i = 1,2)$ into the objective V gives:

$$BV = \frac{\lambda}{\lambda + (1-\lambda)F(s)}(q_1 BY_1 - q_1 BY_2 + q_1 C_1 - q_1 C_2) + \lambda BY_2 + \lambda C_2.$$

Substituting out $q_1(BY_1 - BY_2)$ using the indifference condition obtained earlier, BV becomes:

$$\frac{\lambda}{\lambda + (1-\lambda)F(s)}[\delta(\lambda + (1-\lambda)F(s))s - q_1 s + q_1 C_1] + \lambda BY_2 + \lambda C_2.$$

This is equivalent to maximizing:

$$\begin{aligned}
&(\delta - p)s + pC_1 + BY_2 + C_2 \\
=&(\delta - p)s + pC_1 + C_2 + BQ - (\delta - p)(\lambda + (1-\lambda)F(s))s - p(\lambda + (1-\lambda)F(C_1))C_1 - (\lambda + (1-\lambda)F(C_2))C_2 \\
=&(1-\lambda)[(\delta - p)(1 - F(s))s + p(1 - F(C_1))C_1 + (1 - F(C_2))C_2] + BQ.
\end{aligned}$$

Therefore, the optimal policy in case (1) solves:

$$\begin{aligned}
\max_{s,C_1,C_2} \quad & (\delta - p)\theta(s) + p\theta(C_1) + \theta(C_2) \\
\text{subject to} \quad & BQ - (\delta - p)\Theta(s) - p\Theta(C_1) - \Theta(C_2) \geq 0, \\
& s - C_1 \geq 0, \\
& C_2 - s \geq 0.
\end{aligned}$$

Consider a policy with $s > \hat{S}$. Lowering s weakly increases the objective function because $\theta(\cdot)$ is decreasing for $s > \hat{S}$. Lowering s strictly relaxes the feasibility constraint because $\Theta(\cdot)$ is increasing. Therefore, $s > \hat{S}$ cannot be optimal. Similarly, $C_i > \hat{S}$ cannot be optimal for $i = 1,2$.

Consider any feasible policy that satisfies all the constraints with $(C_1, C_2, s) \in (-\infty, \hat{S}]^3$ not all equal. Then, there is an alternative policy (C_1', C_2', s') such that

$$C_1' = C_2' = s' = \frac{\delta - p}{1+\delta}s + p\frac{p}{1+\delta}C_1 + \frac{1}{1+\delta}C_2.$$

If the original policy satisfies the feasibility constraint, the alternative policy also satisfies the feasibility constraint because $\Theta(\cdot)$ is convex. It strictly increases the objective function because $\theta(\cdot)$ is strictly concave on $(0, \hat{S}]$, $C_2 > 0$ and $s > 0$. It is therefore without loss of generality to only consider policies with $C_1 = C_2 = s = \hat{s}$. For such policies, the problem reduces further to:

$$\max_{\hat{s}\in[0,\hat{S}]} \quad (1+\delta)\theta(\hat{s})$$
$$\text{subject to} \quad BQ-(1+\delta)(s-(1-\lambda)\theta(\hat{s}))\geq 0.$$

The solution to the problem is $\hat{s}=\hat{S}$ if $\hat{S}<s^*$ and $\hat{s}=s^*$ otherwise. The policy corresponding to $C_1=C_2=s=\hat{s}$ is $Y_2=Q-\frac{1+\delta}{B}(\lambda+(1-\lambda)F(\hat{s}))\hat{s}$, $X_1=Y_2+\frac{\hat{s}}{B}+\frac{\delta(Q-Y_2)}{(1+\delta)q_1}$, and $X_2=Y_2+\frac{\hat{s}}{B}$. Because no one gets a low score in School 1 under the optimal policy, the value of Y_1 is not uniquely pinned down as long as all types in School 1 below $\hat{s}$ weakly prefer to get tutoring at stage two. This requires $Y_1\in\left[0,X_1-\frac{\hat{s}}{B}\right]$.

The analyses of the other three cases are similar to that of case (1), except that the weights attached to $\theta(\cdot)$ (in the objective function) and $\Theta(\cdot)$ (in the feasibility constraint) are different. For case (2), the objective function is $\delta\theta(s)+p\theta(C_1)+(1-p)\theta(C_2)$. In this case, it is optimal to set $C_1=C_2=s$. For case (3), the objective function is $\theta(C_2)+\delta\theta(s)$. In this case it is optimal to set $C_2=s$. For case (4), the objective function is $(1-p)\theta(C_2)+(\delta+p)\theta(s)$. In this case it is optimal to set $C_2=s$.

Non-cutoff policies. We consider next policies with the non-cutoff nature, with indifferent types $\underline{s}<C_1<s<C_2<\bar{s}$. By their definitions,

$$\begin{aligned}
C_1 &= B(X_1-Y_1)\\
C_2 &= B(X_2-Y_2)\\
\delta\underline{s} &= p(BX_1-BX_2)\\
\delta s &= p(BY_1-BX_2+s)\\
\delta\bar{s} &= p(BY_1-BY_2)
\end{aligned}$$

where $p=q_1/(\lambda+\lambda\tilde{F})$, with $\tilde{F}=F(\underline{s})+F(\bar{s})-F(s)$.

Following the same steps as before, the feasibility condition can be written as:

$$BY_2+\delta(\lambda+(1-\lambda)\tilde{F})s+C_2-q_1s+q_1(\lambda_1+(1-\lambda_1)G_1(C_1))C_1-(1-q_1)(1-\lambda_2)(1-G_2(C_2))C_2=BQ.$$

We have,

$$
\begin{aligned}
G_1(C_1) &= \frac{F(\underline{s})}{\tilde{F}},\\
G_2(C_2) &= \frac{F(C_2) - pF(\underline{s}) - p(F(C_2) - F(s))}{1 - p\tilde{F}},\\
q_1(\lambda_1 + (1-\lambda_1)G_1(C_1)) &= p(\lambda + (1-\lambda)F(\underline{s})),\\
(1-q_1)(1-\lambda_2)(1-G_2(C_2)) &= (1-\lambda)(1 - pF(\bar{s}) - (1-p)F(C_2)).
\end{aligned}
$$

Plugging these into the feasibility condition, we obtain:

$$BY_2 - (p-\delta)(\lambda + (1-\lambda)\tilde{F})s + p(\lambda + (1-\lambda)F(\underline{s}))C_1 + p(\lambda + (1-\lambda)F(\bar{s}))C_2 + (1-p)(\lambda + (1-\lambda)F(C_2))C_2 = BQ.$$

From the definition of $\underline{s}$, s, $\bar{s}$, C_1 and C_2, we have:

$$
\begin{aligned}
C_1 &= \frac{\delta}{p}\underline{s} + \frac{p-\delta}{p}s,\\
C_2 &= \frac{\delta}{p}\bar{s} + \frac{p-\delta}{p}s.
\end{aligned}
$$

Substitute out C_1 and C_2 in all but the last term in the feasibility condition to get,

$$BY_2 + \delta(\lambda + (1-\lambda)F(\underline{s}))\underline{s} + \delta(\lambda + (1-\lambda)F(\bar{s}))\bar{s} + (p-\delta)(\lambda + (1-\lambda)F(s))s + (1-p)(\lambda + (1-\lambda)F(C_2))C_2 = BQ.$$

Therefore, the condition $Y_2 \geq 0$ is equivalent to:

$$BQ - \delta\Theta(\underline{s}) - \delta\Theta(\bar{s}) - (p-\delta)\Theta(s) - (1-p)\Theta(C_2) \geq 0.$$

Similarly, plugging the indifference conditions for C_1, C_2 and s into the objective function gives:

$$BV = \frac{\lambda}{\lambda + (1-\lambda)\tilde{F}}\left[\delta(\lambda + (1-\lambda)\tilde{F})s - q_1 s + q_1 C_1\right] + \lambda BY_2 + \lambda C_2.$$

This is equivalent to maximizing:

$$
\begin{aligned}
&-(p-\delta)s + pC_1 + C_2 + BY_2\\
=&BQ + (1-\lambda)\left[-(p-\delta)(1-\tilde{F})s + p(1-F(\underline{s}))C_1 + p(1-F(\bar{s}))C_2 + (1-p)(1-F(C_2))C_2\right].
\end{aligned}
$$

Substituting out C_1 and C_2 in all but the last term from the above expression, the problem reduces to:

$$\begin{aligned}
\max_{\underline{s},s,\bar{s},C_2} \quad & \delta\theta(\underline{s})+\delta\theta(\bar{s})+(p-\delta)\theta(s)+(1-p)\theta(C_2) \\
\text{subject to} \quad & C_2=\frac{\delta}{p}\bar{s}+\frac{p-\delta}{p}s, \\
& BQ-\delta\Theta(\underline{s})-\delta\Theta(\bar{s})-(p-\delta)\Theta(s)-(1-p)\Theta(C_2)\geq 0, \\
& s-\underline{s}\geq 0, \\
& \bar{s}-s\geq 0.
\end{aligned}$$

The same argument as before shows that we only need to consider policies with $(\underline{s}, s, \bar{s}, C_2) \in (-\infty, \hat{S}]^4$, where at least $\bar{s} > 0$. Furthermore, Jensen's inequality implies that any such policy is dominated by a policy with $\underline{s} = s = \bar{s} = C_2$. So non-cutoff policies are never optimal.

References

Ball, Ian (2021). "Scoring Strategic Agents." Working paper, Yale University.

Bray, Mark (1999). *The Shadow Education System: Private Tutoring and its Implications for Planners.* Paris: International Institute for Educational Planning.

Bray, Mark and Percy Kwok (2003). "Demand for Private Supplementary Tutoring: Conceptual Considerations, and Socio-economic Patterns in Hong Kong." *Economics of Education Review* 22: 611–620.

Chan, Jimmy and Erik Eyster (2003). "Does Banning Affirmative Action Lower College Student Quality?" *American Economic Review* 93(3): 858–872.

Chan, Jimmy, Xian Fang, Zhi Wang, Xianhua Zai, and Qinghua Zhang (2020). "Valuing Primary Schools in Urban China." *Journal of Urban Economics* 115: 103183.

Chan, William, Hao Li, and Wing Suen (2007). "A Signaling Theory of Grade Inflation." *International Economic Review* 48(3): 1065–1090.

Coate, Stephen and Glenn C. Loury (1993). "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review* 83(5): 1220–1240.

Curry, Philip A. and Steve Mongrain (2009). "Deterrence in Rank-Order Tournaments." *Review of Law and Economics* 5(1): 723–740.

Ederer, Florian, Richard Holden, and Margaret Meyer (2018). "Gaming and Strategic Opacity in Incentive Provision." *RAND Journal of Economics* 49(4): 819–854.

Fang, Dawei and Thomas Noe (2019). "Less Competition, More Meritocracy?" Working paper, University of Gothenburg.

Farmer, Amy and Dek Terrell (2001). "Crime versus Justice: Is There a Trade-Off?" *Journal of Law and Economics* 44(2): 345–366.

Frankel, Alex, and Navin Kartik (2019). "Muddled Information." *Journal of Political Economy* 127(4): 1739–1776.

Frankel, Alex, and Navin Kartik (2021). "Improving Information from Manipulable Data." Working paper, University of Chicago.

Goldman, Eitan, and Steve L. Slezak (2006). "An Equilibrium Model of Incentive Contracts in the Presence of Information Manipulation." *Journal of Financial Economics* 80(3): 603–626.

Gilpatric, Scott M (2011). "Cheating in Contests." *Economic Inquiry* 49(4): 1042–1053.

Gilpatric, Scott M. and Cristina M. Reiser (2017). "Why Zero Tolerance of Misconduct is Undesirable in Contests." *Economic Inquiry* 55(2): 1145–1160.

Hardt, M., Megiddo, N., Papadimitriou, C. and Wootters, M., (2016) "Strategic Classification." *Proceedings of the 2016 ACM conference on innovations in theoretical computer science*: 111–122.

Li, Jin, Arijit Mukherjee, and Luis Vasconcelos (2021). "Learning to Game the System." *Review of Economic Studies,* 88(4): 2014–2041.

Lui, Hon-Kwong and Wing Suen (2005). "The Shrinking Earnings Premium for University Graduates in Hong Kong: The Effect of Quantity or Quality?" *Contemporary Economic Policy* 23(2): 242–254.

Kartik, Navin (2009). "Strategic Communication with Lying Costs." *Review of Economic Studies* 76(4): 1359–1395.

Kartik, Navin, Marco Ottaviani, and Francesco Squintani (2007). "Credulity, Lies, and Costly Talk." *Journal of Economic Theory* 134: 93–116.

Korn, Melissa and Jennifer Levitz (2020). *Unacceptable: Privilege, Deceit and the Making of the College Admissions Scandal.* New York: Portfolio.

Lazear, Edward P. (2006). "Speeding, Terrorism, and Teaching to the Test." *Quarterly Journal of Economics* 121(3): 1029–1061.

Lee, Chong-Jae, Hyun-Jeong Park, and Heesook Lee (2009). "Shadow Education Systems." Pp. 901–919 in *Handbook of Education Policy Research.* London: Routledge.

Lundberg, Shelly J. and Richard Startz (1983). "Private Discrimination and Social Intervention in Competitive Labor Market." *American Economic Review* 73(3): 340–347.

Milgrom, Paul (1988). "Employment Contracts, Influence Activities, and Efficient Organization Design." *Journal of Political Economy* 96(1): 42–60.

Milgrom, Paul and John Roberts (1988). "An Economic Approach to Influence Activities in Organizations." *American Journal of Sociology* 94: S154–S179.

Nitzan, Shmuel (1994). "Modelling Rent-Seeking Contests." *European Journal of Political Economy* 10(1): 41–60.

Park, Julie J., and Sooji Kim (2020). "Harvard's Personal Rating: The Impact of Private High School Attendance." *Asian American Policy Review* 30: 2–80.

Popham, W. James (2001). "Teaching to the Test?" *Educational Leadership* 58(6): 16–21.

Prendergast, Canice and Robert H. Topel (1996). "Favoritism in Organizations." *Journal of Political Economy* 104(5): 958–978.

Ramey, Garey and Valerie Ramey (2009). "The Rug Rate Race." NBER Working Paper 15284.

Siegel, Ron (2009). "All Pay Contests." *Econometrica* 77(1): 71–92.

Tullock, Gordon (2001). *Efficient Rent-Seeking.* Boston: Springer.

Volante, Louis (2004). "Teaching to the Test: What Every Educator and Policy-Maker Should Know." *Canadian Journal of Educational Administration and Policy* 35: 1–6.

Whitmeyer, Mark (2019). "Opacity Design in Signaling Games." Working paper, University of Bonn.