An Economic Analysis on the Exam Aspects of the Exam Hell

1. Introduction

In some East Asian countries like Japan, Korea, and Taiwan, students live in the "exam hell", a term often used to describe the miserable lives of those students struggling to survive under the shadow of the college entrance exam. The facts, the causes, and the consequences of the exam hell have attracted the attention of scholars from many academic disciplines.[1] This paper focuses on one of the possible causes of the exam hell, or more specifically, one of the possible causes of the phenomenon that students spend too much time and study too hard for the college entrance exam. Many attribute the causes of this over-studying problem to some cultural and economic-development factors such as Confucianism, the tradition of governmental official exam, and the dual-economy problem. However, we think more attention should be paid on how much the over-studying problem is related to the designs of the exam system. This is because the exam system, if it is (at least partially) responsible for the exam hell, could be adjusted more rapidly and easily to mitigate the over-studying problem than those cultural and economical causes. Economic theories developed to study the effort-inducing function and the selection function of institutions are well suited for studying this issue, since these two functions are the ones that the college entrance exam is designed to serve. It is the purpose of this paper to apply these economic theories to study the relationship between the designs of the exam and the over-studying problem.

According to the educational measurement literature, the college entrance exam in East Asia can be categorized as a "norm-referenced" "achievement" test.[2] The exam is an "achievement" test because it is designed to evaluate students' understanding on certain materials which are strictly confined within the high school curriculum. And the exam is "norm-referenced" because the admission decisions are based on and only on students' rankings, that is, based only on students' relative performance. Many countries use exams of different designs for their college admission process. The alternative to the achievement test is the "aptitude" test that evaluates students' intelligence instead of their learning achievements. The SAT test in the U.S. is considered by many as an aptitude test. The alternative to the norm-referenced test is the "criterion-referenced" test that evaluates students' performances by comparing them to a standard defined on how well students learn certain material. In some European countries, whether students are admitted to (public) colleges is determined by whether they pass the exam held at the end of their secondary education

[1] The following papers are examples of the researches including this term in the title. Lee (2003), Lee & Larson (1999), and Holman (1991) considered the distress and depression caused by long hours of studying. Berwick & Ross (1989) studied the impacts of exam hell on students' learning motivation and attitude. Ogura (1987) addressed general educational problems. Also see Appendix A for the evidences of Korean high-school students' long studying hours documented in several different sources.

[2] See Sax (1989), Nunnally (1964), and Kubiszyn & Borich (2003) for the background of the literature.

(i.e., whether they meet the criterion).[3] The main concern of this paper is whether or not and, if so, to what extent these two features are responsible for the over-studying problem, and what institutional changes can be made to mitigate the problem.

Our hunch that the norm-referenced feature may be responsible for the over-studying problem is best illustrated by what is described in Ellinger and Beckham (1997).

> *"[Korean high school] students are accustomed to getting on by no more than four hours of sleep, preceded by periodic visits from their mothers, who bring them coffee and remind them that, while they sleep, others are studying."*

Becker and Rosen (1992) also implied that they think that the norm-referenced test tends to induce more effort than the criterion-referenced one:

> *"Why do Japanese parents believe that their children must attend cram school? It may not be a high national standard. It may be that they realize that such action is needed to get an edge on other students."*

Students and their parents would not have to worry about how much other students study if the exam is a criterion-referenced one. In that case, all they need to worry about is to beat the criterion, not to beat other students.

Examples about the relationship between the achievement test and the exam hell can also be found in the literature. Noah and Eckstein (1989b) documented the different effort-inducing performances between an achievement test (the Joint First-stage Achievement Test (JFSAT) in Japan) and an aptitude test (the SAT in the U.S.) Other than this distinction, these two exams are very similar:

> *"Of the eight countries in our study, Japan and the United States are the only two to have adopted a virtually exclusive multiple-choice, machine-scorable format for the university entrance examinations."*

However, their effort-inducing outcomes are dramatically different:

> *"[In Japan], the competition is so intense and the pressures are so great during the secondary school period that the universities complain that students arrive burned-out, determined to make up for their lost youth, and unwilling to continue to study hard. [While, in the United States] the complaint is about the lack of challenge that many high schools offer to their students, and the shock that college freshmen can receive when confronted for the first time with major demands upon their time and intellect."*

[3] Examples for such exams are Germany's Abitur, France's Baccalaureat, and Austria's Matura. For more on cross-country introductions and comparisons of the exam systems related to college admission, see Stevenson & Lee (1997), Noah & Eckstein (1989a, b), and Feuer & Fulton (1994).

The effort-inducing function and the selection function of educational measurements in general, and college entrance exams in particular, have been studied theoretically by economists. Both Stiglitz (1975) and Costrell (1994) studied criterion-referenced exams. Stiglitz assumed that the only thing an educational institution does is to measure and label students' ability, and that students' future income depends on the ability they are labeled. He also assumed a positive relationship between the educational criterion set by an educational institution and how precise the institution measures students' ability. He used this relationship to derive the preferences of students of different abilities on different levels of criterion, and he studied how these different preferences generate different levels of academic criterion through different political processes. In short, Stiglitz explored the political economy of the selection function of the educational standard. Costrell (1994) also studied educational standard, but he focused more on its effort-inducing function. He assumed the school sets one and only one educational standard (a criterion) for all students who may have different ability levels. The students meeting the criterion are awarded a degree informing their future employers that they had accomplished the level of learning success defined by the criterion. Thus, as in Stiglitz's model, different students prefer different criteria and different criteria are chosen through different political processes. Becker and Rosen (1992) studied both criterion-referenced tests and norm-referenced tests. Similar to Stiglitz and Costrell, they focused on the study of welfare and the political economy involved in the process of determining the criterion. The forms of the criterion-referenced and the norm-referenced exams considered by them are fairly simple and the effort-inducing and selection functions of these two kinds of exams are not thoroughly compared in their paper. This paper intends to complement their work by exploring in more detail the difference in inducing effort and selecting students between the criterion- and norm-referenced tests.

Economists have been studied the superiority between the relative-performance rewarding mechanism and the absolute-performance rewarding mechanism. Our effort in comparing the norm-referenced and the criterion-referenced tests can be thought as an extension of this study, as the norm-referenced system relies on students' relative performance and the criterion-referenced system relies on students' absolute performance to make admission decisions. Started by Lazear and Rosen (1981) and followed by Green and Stokey (1982), Nalebuff and Stiglitz (1983), O'Keeffe, Viscusi, and Zeckhauser (1984), and Knoeber (1989), much research had been devoted to explain the popularity of relative-performance mechanism in the economy and to delineate the conditions under which the relative-performance mechanism performs better than the absolute-performance mechanism. Our paper differs from the literature in that the comparison we conduct is based on a slightly different benchmark. The literature allows the authority (the employer, or the contest-designer) to manipulate the prize-structure (the number and size of the prizes) of the rewarding mechanisms to maximize profit. In contrast, we let the two systems operate under exactly the same prize-structure: the same number of students competing for the same number of college positions. And we assume the exam authority cannot change this prize structure. The literature concludes that the relative-performance mechanism is better than the absolute-performance mechanism in two situations: when there are common shocks involved in competitors' measured performance (which would be filtered out if performance is measured relatively) and when the measurement cost of relative performance is much lower than that of the absolute performance. Our research finds some new aspects, besides these two reasons, that the relative-performance mechanism is superior to the absolute-performance mechanism. In our model, agents are all assumed to be risk neutral (so the problem of common shocks is not relevant) and the two rewarding mechanisms are assumed to have the same measurement cost. We obtain the result that under certain circumstances the relative-performance mechanism is more flexible in inducing effort than the absolute-performance mechanism and in this sense the former is superior to the latter.

The difference between the aptitude test and the achievement test is, to our knowledge, hardly studied by economists. Harbaugh (2001) took it as given that the achievement test should be used to motivate effort and the aptitude test to select students. Based on this point, together with his assumption that colleges prefer high-aptitude-low-effort students to low-aptitude-high-effort students, he claimed and empirically tested the proposition that those American colleges facing stronger demand tend to put more weights on the aptitude test (SAT) than on other achievement tests[4] in their admission process.

The major part of our work in this paper is devoted to the comparison of the norm- and criterion-referenced tests. We follow economics terminology to call the norm-referenced test the relative-performance system (the RP system) and call the criterion-referenced test the absolute-performance system (the AP system). After that we modify our model to study the difference between using an achievement test and using an aptitude test. In the end we apply our model to study two other related issues: to compare the selecting function of different exam systems, and to discuss the relationship between students' effort level and the admission rate.

The first result we obtain is that, under the same admission rate, the AP system in general induces more effort than the RP system. And the RP system may induce more effort only when students are substantially different and they are not well informed about the ability of their opponents. Namely, the competition among students alone should not be blamed for the over-studying problem. It is the competition under uncertainty that may be responsible for the problem. The second result is that the effort-inducing function of the RP system is sensitive to the size (the number of students) and the informational structure of the

[4] Examples for such achievement tests are ACT and class-rankings from high school.

competition pools. Under the RP systems, students' effort can be effectively lowered, in case it is desired, by downsizing the competition pools (without changing the admission rate and the distribution of students' ability) and at the same time informing students the ability levels of their opponents. The third result, following directly from the second one, is that the exam authority can manipulate the size and the informational structure under the RP system to induce a wide range of effort that, in some situations, includes the level induced by the corresponding AP system. This flexibility makes the norm-referenced test (a relative-performance mechanism) superior to the criterion-referenced test (an absolute-performance mechanism). The fourth result concerns the selection function. When the number of students contending in the pool is small and students are similar in ability, the AP system selects better in the sense that it gives high-ability students a relatively higher chance to enter college. The RP system may select better when the number of students is small, students are fairly different in ability, and students do not know the ability of their opponents. The fifth result is that adding some aptitude-test-type questions to an achievement test lowers students' effort and at the same time improves the selection function. The last result is that creating more college positions (and thus increasing the admission rate) may not release the stress from students as many think. Doing so may actually make students study more.

We investigate the difference between the AP and the RP systems in the next section. In section 3 we study the distinction between the achievement test and the aptitude test. Two applications of our model are presented in section 4. And section 5 concludes our study.

2. The AP system versus the RP system

In order to isolate the differences caused by these two systems from other factors, in this section we assume that both of the systems operate with a pure achievement test. We will modify our model to study the difference between the aptitude test and the achievement test in section 3.

2.1.1. The General Setup

Consider a certain number of university positions, n, are to be awarded somehow to a certain number of students, N ($N > n$). Whether or not a student is given a position is determined entirely by a once-and-for-all measure of his academic achievement: his score in the college entrance exam. The score of a student is a noisy measure of his learning accomplishment. Students have no control over the noisy factor, but they can increase the expected value of their scores (and thus their chances of winning a position) by investing more effort to prepare for the exam. The two exam systems generate different results simply because they use students' scores in different ways to make the admission decisions.

For simplicity, we assume throughout this paper that all of the college positions are identical and we use V to denote this common value of a college position.

Students only differ in their intelligence (henceforth their "ability"). We first consider the case that all students are identical (the homogeneous case), and later we allow their abilities to be different (the heterogeneous case). The homogeneous case allows us to concentrate on the effort-inducing function of the exam systems. The heterogeneous case not only allows us to study the selection function of the exam, but also brings us different results on the effort-inducing function of the exams. We assume students' ability only affects their cost of preparing for the exam: in order to achieve the same academic achievement, low-ability students need to pay higher cost than high-ability students.[5]

In the heterogeneous cases, students' knowledge about their and others' ability levels becomes relevant. Throughout this paper, we assume students always know their own ability and the exam authority does not observe any student's ability.[6] For the information on other students' ability, we will consider two cases. In the first part of the heterogeneous case we consider the situation under which all students' ability is common knowledge to every student in the competition pool. And in the second part we consider the situation where students only know the distribution of the combination of their competitors' ability. We use A_i to denote the ability level of student i.

In the real world, no exam measures students' academic achievement perfectly. We model this measurement error using the method Lazear and Rosen (1982) adopted in their study on the rank-order tournaments. That is, we assume every student's score to be a random variable and the effort he invests only affects the mean of the random variables.[7] We denote student i's score as y_i. And we assume $y_i = \mu_i + \varepsilon_i$, where μ_i is student i's "effective effort" and ε_i is the measurement error associated with the exam.[8,9] The measurement

[5] We do not allow students' ability to affect their scores directly until section 3. Some papers in the rent-seeking literature consider the case that different competitors have different tastes and thus value the same prize differently. We do not consider that possibility in this paper, but our model can accommodate it easily.
[6] The authority is assumed to know the distribution of the composition of students' ability only.
[7] The RP system in our model is similar to the contest games studied by the rent-seeking literature. More precisely, the RP system in our model can be considered as the contest games commonly used in that literature with a different form of contest success function. For more about the contest function used in the rent-seeking literature, see Appendix B.
[8] We emphasize the term "effective" to distinguish "effort" from "time". It is the effort that can be efficiently translated into score that counts. Henceforth, we will drop the term effective and use "effort" for simplicity.

error is assumed to be independent to students' effort and ability, namely, to be identical across all students. The above assumptions imply that students with different abilities would have the same score distribution if they invest the same level of effort. We further assume that the distribution of the measurement error belongs to the logistic distribution.[10] In short, we assume that student i's score is a random variable characterized by a logistic distribution with mean μ_i. And we use $F_i(\cdot)$ and $f_i(\cdot)$ to denote the cumulative distribution function (cdf) and the probability density function (pdf) of the distribution.

We assume all students are rational investors/gamblers. They face an investment adventure that the winners receive a prize V and the losers get nothing. Since there are some uncertain factors involved in the process of determining the winners, students have no full control over their success. What students surely know is that investing more effort can increase their chances of winning a prize. We assume all students are risk neutral. Thus, when student i invests μ_i, his "profit" function can be written as $\pi = V \cdot P(\mu_i \mid \cdot) - C(\mu_i \mid \cdot)$, where $P(\mu_i \mid \cdot)$ is his chance of winning a position and $C(\mu_i \mid \cdot)$ is the cost he incurs. The problem that student i faces in this model is to find the effort level that maximizes his profit.

To make this problem tractable and meaningful, we assume that this adventure is worth investing at the beginning: the benefit produced by the first unit of effort (the marginal benefit of the first unit) is greater than the cost of it (the marginal cost of the first unit). We also assume that, as the effort invested increases, the marginal cost keeps increasing and the marginal benefit is either decreasing or increasing at a relatively lower rate. Therefore, beyond some level of effort the marginal cost will become greater than the marginal benefit and no effort will be worth investing after that. The optimal level is where marginal benefit equals the marginal cost.[11] The condition that the marginal benefit equals marginal cost is nothing but the first-order condition of the individual's optimization problem, which is simply $V \cdot P'(\mu_i \mid \cdot) = C'(\mu_i \mid \cdot)$.

As mentioned above, we assume increasing marginal cost of investing effort. This assumption can be easily justified. Recall that, in this paper, by "effort" we mean "effective effort": it is the effort that can be

---

[9] There are many possible sources of this measurement error: students guess on multiple choice questions, graders are neither purely objective nor mistake-free when they mark essay questions, some questions may be poorly written and/or beyond the scope of the curriculum, some students may not be physically or mentally sound on the exam day, and so on. In short, the measurement error captures all the uncertain factors that occur after students invest their effort and before their scores come out.

[10] The logistic distribution is like the normal distribution: it is symmetric around the mean and has a bell shape. We use the logistic distribution instead of the normal distribution because its cdf is simple and it allows us to formally prove our results. We can show that all of our results also hold for the normal distribution by using numerical methods.

[11] In economics terminology, we assume both the participation conditions and the second order conditions are satisfied, and thus we can concentrate our analysis on the first order conditions. The satisfactions and possible violations of the participation conditions and the second order conditions are discussed in Appendix B.

translated productively into score. Using this explanation of effort, we offer two reasons to support this increasing-marginal-cost assumption. On the one hand, students become tired and less productive as study goes on and thus it takes longer time for students to carry out the same unit of effective effort. On the other hand, students suffer more from each hour's studying as the studying time gets longer. Throughout this paper, we employ a simple function to capture this property: we assume the cost function $C(\mu_i \mid \cdot)$ to be $C(\mu_i \mid \cdot) = \mu_i^2 / 2A_i$. One can see that this function is consistent with our assumption as the marginal cost function $C'(\mu_i \mid \cdot) = dC(\mu_i \mid \cdot) / d\mu_i = \mu_i / A_i$ is an increasing function of μ_i. This cost function also leads to the outcome that, when investing the same amount of effort, high-ability students face both a lower total cost and a lower marginal cost.

The marginal-benefit part is more complicated since it differs among different exam systems. The marginal benefit consists of two parts: V and $P'(\mu_i \mid \cdot)$. V is the value of a college position that is common to every student, and $P'(\mu_i \mid \cdot)$ is the marginal product of student i's effort on his winning probability: the increase in his chance of winning brought by an infinitesimal amount of increase in his effort. This term is important for most of our analyses in this paper, and we use "MPWP" (the marginal product on winning probability) to denote this term henceforth. The two exam systems generate different forms of MPWP, and we use the next two subsections to study them separately.

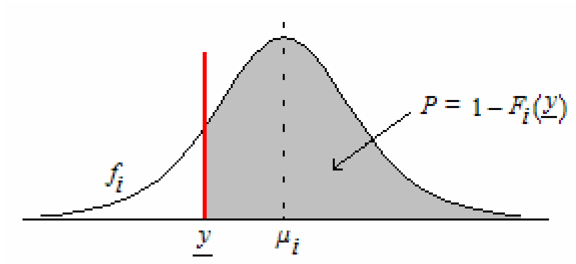
2.1.2. The MPWP under the AP system

We consider a very simple form of the AP system: the educational authority sets a criterion and admits every single student whose score is greater or equal to the criterion.[12] We assume the criterion is set before students invest effort.

We need to first consider the relationship between the effort invested by a student and his chance of being admitted under this exam system. One can see this relationship clearly from < Figure 1 >. Suppose that student i invests an effort level μ_i and gets a score distribution characterized by $f_i(\cdot)$. If the authority sets the criterion equal to $\underline{y}$, then student i's chance of being admitted would be the portion of his score

[12] Because of our assumption that the score is a random variable belonging to the logistic distribution and the fact that the domain of the logistic distribution is from negative infinity to positive infinity, no student can be sure that he will be admitted no matter how hard he works. The question would be less interesting without this random factor: students will either invest to meet the criterion or quit.

distribution greater or equal to $\underline{y}$. In < Figure 1 >, this chance is shown as the shaded area: the area below the pdf and on the right hand side of $\underline{y}$. In other words, the probability that student i is admitted when he faces a criterion $\underline{y}$ and invests μ_i is $1-F_i\left(\underline{y}\right)$.



< Figure 1>

The MPWP of student i's effort can be derived easily from < Figure 1 >. If student i is to increases his effort by a small amount $\Delta\mu$ from μ_i, then his score distribution would be shifted to the right by the same magnitude and the shaded area would increase roughly by $\Delta\mu \cdot f_i\left(\underline{y}\right)$. The MPWP can be approximated as $\Delta\mu \cdot f_i\left(\underline{y}\right) / \Delta\mu = f_i\left(\underline{y}\right)$. That is, the MPWP of student i's effort is equal to the value of the pdf of the score distribution measured at the criterion.[13]

To briefly summarize, under the AP system, we can write the marginal benefit of student i's effort as $V \cdot f_i\left(\underline{y}\right)$. Together with the marginal cost function $\left(\mu_i / A_i\right)$, the condition for student i's optimization problem is simply $V \cdot f_i\left(\underline{y}\right) = \mu_i / A_i$.

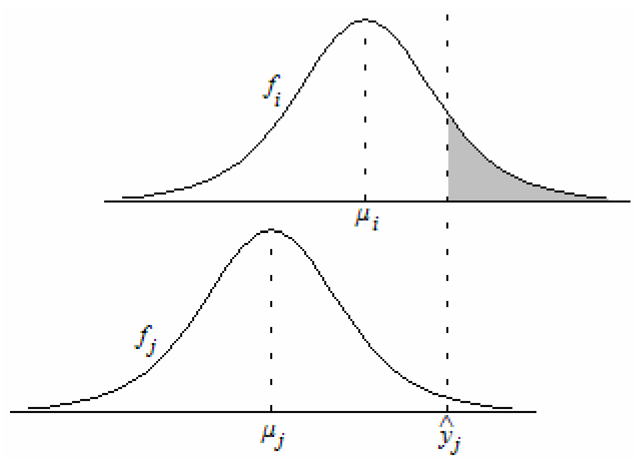
[13] To show this result formally, we first define the pdf and cdf of the measurement error ε as $\theta(\cdot)$ and $\Theta(\cdot)$ respectively. Since $\theta(\cdot)$ and $f_i(\cdot)$ present two random variables that are identical except having different means, this property $f_i(\hat{x}) = \theta(\hat{x} - \mu_i)$ must hold for all $\hat{x}$. This in turn means $F_i(\hat{x}) = \Theta(\hat{x} - \mu_i)$. To find the MPWP mathematically, we differentiate the winning probability $1-F_i\left(\underline{y}\right)$ with respect to μ_i. And the result is

$$\partial\left[1-F_i\left(\underline{y}\right)\right] / \partial\mu_i = -\partial F_i\left(\underline{y}\right) / \partial\mu_i = -\partial\Theta\left(\underline{y}-\mu_i\right) / \partial\mu_i = \theta\left(\underline{y}-\mu_i\right) = f_i\left(\underline{y}\right).$$

2.1.3. The MPWP under the RP system

Under the RP system, the exam authority ranks students from the highest score to the lowest and admitted students into college from the top of the ranking till all positions are taken.[14] That is, instead of facing a given criterion, students under the RP system face and compete with other students.

The method of deriving the MPWP of students' effort under this kind of exam system can be illustrated by a 2-student example. Consider the case there are two students, i and j, competing for one position (hereafter the 1-for-2 case). We assume that students "summit" their effort levels to the exam authority at the same time without observing the other student's effort, and then the authority compares the scores it observed to determine the winner. We assume μ_i and μ_j to be their effort levels and $f_i(\cdot)$ and $f_j(\cdot)$ to be the pdfs of their score distributions, as illustrated in < Figure 2 >.



< Figure 2 >

What is the MPWP of student i's effort in the example? Firstly, he does not know in advance what student j's score will be. If student j's score happens to be $\hat{y}_j$, then we could think $\hat{y}_j$ as the criterion student i needs to beat. And from our analysis on the AP system we know the MPWP of student i's effort would simply be $f_i(\hat{y}_j)$: an increase in μ_i shifts student i's score distribution to the right and increases the area representing his winning probability, the shaded area, by $f_i(\hat{y}_j)$. This analysis can be applied to all possible values of student j's score. Thus the MPWP of student i's effort is the average of all possible values of

[14] In this paper, we consider the cases that the authority has the power to change the size of the competition pool (the number of students in a pool) given the condition that all groups have the same admission rate.

$f_i\left(\hat{y}_j\right)$ weighted by $f_j\left(\hat{y}_j\right)$, the chance that student j's realized score happens to be $\hat{y}_j$. In other words, the MPWP of student i's effort in this RP system is $\int\limits_{\hat{y}=-\infty}^{\infty} f_i(\hat{y}) f_j(\hat{y}) d\hat{y}$.[15]

This result that under the RP system the MPWP of a student's effort is the average of the pdf of his score distribution weighted by the pdf of his opponent's score distribution is very important to our analyses. And this result can be easily extended to the more complicated cases. The key for this extension is that all of the more complicated cases, with more than one prize and/or with more than two players, can always be reduced to a 1-for-2 case, the case that one would get a prize if he beat the opponent and would get nothing if he gets beat. For example, in a 2-for-10 competition, the only thing matters for one of the competitors is whether or not he beats the second best of the other nine students. If he does, he would get a prize and he would not do any better by beating one more. If he does not, he would get nothing and he would not do any worse by losing to more competitors.[16] After identifying this "imaginary opponent", we can apply the method of order statistics to derive his score distribution. Throughout this paper, we will transform every RP competition into a 1-for-2 competition between the student in question, student i, and his imaginary opponent. And we use $g(\cdot)$ and $G(\cdot)$ to denote the pdf and cdf of the score distribution of this imaginary opponent. Therefore the MPWP of student i's effort can always be written as $\int\limits_{\hat{y}=-\infty}^{\infty} f_i(\hat{y}) g(\hat{y}) d\hat{y}$, the average of the pdf of student i's score distribution using the pdf of his opponent's score distribution as the weights. In the remaining of this paper, we will simply apply this result to derive the MPWPs under all of the RP systems and skip formal derivations.

---

[15] See Appendix C for formal derivation of this result. In the tournaments literature the MPWP is usually presented in the "traditional" way: $\int\limits_{\hat{\varepsilon}=-\infty}^{\infty} \theta(\hat{\varepsilon})\theta\left(\mu_i-\mu_j-\hat{\varepsilon}\right)d\hat{\varepsilon}$ where $\theta(\cdot)$ is the pdf of the measurement error ε. (For example, see Gibbons (1992).) This expression is identical to ours: both of them are the integration of the product of two pdfs that represent two distributions which are identical except their means are $\left(\mu_i-\mu_j\right)$ apart from each other. We present the score distributions in this way because it is easier for us to calculate the order statistics. The disadvantage of our approach is that it is more difficult to see the role of the choice variable. We want to attempt to make it clear for our readers: student i chooses effort (μ_i) to affect the gap between his mean score and his opponent's means score. In the "traditional" way, it shows up in the term $\theta\left(\mu_i-\mu_j-\hat{\varepsilon}\right)$ directly. While in our way, a change in μ_i changes the mean of the distribution represented by $f_i(\cdot)$ and thus implicitly changes the gap between the means of the two distributions.

[16] Recall that we assume all prizes are identical for everyone.

2.2.1. Homogeneous students in a competition with a 50% admission rate

After we figure out the MPWPs of students' effort under the AP and the RP systems, we now move on to compare the effort induced by these two exam systems. We would like to start with a benchmark case: some homogeneous students competing under a 50% admission rate. In the homogeneous case, we assume $A_i = A$ for all i.

Under the AP system, the first order condition for student i's optimization problem is $V \cdot f_i\left(\underline{y}\right) = \mu_i / A$. To derive the result for this case, we simply need to add one additional condition: the condition under which every student has exactly a 50% chance of being admitted. Since we assume the score distribution is symmetric around its mean, this 50% condition implies that every student invests his effort to the level such that the mean of his score distribution exactly equals the criterion. That is, it must be the case that $\mu_i = \underline{y}$. In short, if the exam authority in an AP system is to enforce a 50% admission rate, he would have to set a criterion $\underline{y}$ such that all students would find it optimal to choose their effort level (and thus their mean scores) right equal to this criterion.

Putting together the first order condition, $V \cdot f_i\left(\underline{y}\right) = \mu_i / A$, and the 50% condition, $\mu_i = \underline{y}$, the solution for this case is simply

$$V \cdot f_i\left(\mu_i\right) = \mu_i / A \qquad (1)$$

The solution to (1) exists and is unique.[17] Notice that the MPWP of students' effort is $f_i\left(\mu_i\right)$, which is the maximum value of the pdf of the score distribution. Since under this system all students are identical and each of them "competes" with the criterion individually, the equilibrium solution of the whole system is nothing more than applying (1) to all students.

For the RP system, we start with the simplest case: the 1-for-2 case. We call these two students "1" and "2", assume their effort levels to be μ_1 and μ_2, and denote the pdfs of their score distributions as $f_1(\cdot)$ and

[17] The solution for (1) is $\mu_i = A \cdot V \cdot f_i\left(\mu_i\right)$. Since $f_i\left(\mu_i\right)$ is a known number given the pdf of the measurement error and the values of A and V are also given exogenously, the solution exists and is unique.

$f_2(\cdot)$. According to our previous study, the MPWP for student 1's effort is $\int_{\hat{y}=-\infty}^{\infty} f_1(\hat{y}) f_2(\hat{y}) d\hat{y}$ and the MPWP for student 2's effort is $\int_{\hat{y}=-\infty}^{\infty} f_2(\hat{y}) f_1(\hat{y}) d\hat{y}$.

Given μ_2, the optimal level of effort for student 1 can be solved from the first order condition $V \cdot \int_{\hat{y}=-\infty}^{\infty} f_1(\hat{y}) f_2(\hat{y}) d\hat{y} = \mu_1 / A$; and, given μ_1, the condition for student 2's optimal effort is $V \cdot \int_{\hat{y}=-\infty}^{\infty} f_2(\hat{y}) f_1(\hat{y}) d\hat{y} = \mu_2 / A$. Using the concept of Nash equilibrium, the equilibrium level of μ_1 and μ_2 can be obtained by solving the previous two equations simultaneously.

This equilibrium solution can be further simplified. Because student 1 and student 2 are identical, we expect them to choose the same effort level. That is, in the equilibrium we must have $\mu_1 = \mu_2$, and thus $f_1 = f_2$. Let's define $\mu \equiv \mu_1 = \mu_2$ and $f(\cdot) \equiv f_1(\cdot) = f_2(\cdot)$. The equilibrium level of effort, which exists and is unique[18], can be derived simply from

$$V \times \int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) f(\hat{y}) d\hat{y} = \mu / A \qquad (2)$$

With (1) and (2), we can compare the effort levels brought by the AP system and the 1-for-2 RP system.

First of all, we need to explain how to compare these two systems. Under both AP and RP system, students' optimal effort levels are solved from equations with the following form: $V \cdot \text{MPWP} = \mu / A$. They have different solutions only because they have different MPWPs. We claim that if one system generates a greater value of MPWP than the other for all possible values of μ, then the system would induce a higher level of effort than the other. This is very straightforward. Suppose we put one system's equilibrium effort level into the equilibrium condition of the other and find that the marginal benefit, $V \cdot \text{MPWP}$, is greater than the marginal cost, μ / A. This tells us that the second system generates a higher level of MPWP from the first

[18] The solution for (2) is $\mu_i = A \cdot V \cdot \int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) f(\hat{y}) d\hat{y}$. Since $\int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) f(\hat{y}) d\hat{y}$ is a known number given the pdf of the measurement error and the values of A and V are given exogenously, the solution exists and is unique.

system's equilibrium effort. And the equilibrium condition for the second system can only be achieved by increasing the effort level (and thus lowering the gap between marginal cost and marginal benefit). Thus, to compare the effort levels induced by these two systems, we only need to compare the MPWPs they generate from the same value of μ.

The MPWP under the AP system for the 50% admission rate is simply $f(\mu)$, and that under the 1-for-2 RP system is $\int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) f(\hat{y}) d\hat{y}$. It is easy for one to see that the AP system always generates a larger value of MPWP than the RP system: the MPWP under the AP system equals the maximum of the density function while the MPWP under the RP system is an average value of the density function.[19] In short, the AP system under a 50% admission rate induces more effort than the 1-for-2 RP system.

This result is somewhat counterintuitive and deserves more detailed discussions. Under the AP system, students know exactly the marginal product of their effort and know that the marginal product, in this case, happens to be the highest possible pdf value. Under the RP system, the marginal product of students' effort depends on their opponent's realized score. If the opponent's realized score happens to be equal to the mean score of the student in question, then the marginal product of the student would be as high as that under the AP system. But the chance for this to happen is very small. Most of the time the opponent's realized score will be different from the mean score and the marginal product will be lower, in some cases much lower, than that under the AP system. In short, under the condition that they have the same admission rate of 50%, the effort used to try to beat another student is less productive than the effort used in the attempt to beat a fixed criterion, since in the former there is a high probability that some effort is "wasted" in some situations with very low productivity. One can easily see that this "wasted-effort" explanation should apply well to all the cases with admission rates near 50%.

The other explanation regarding this result is that it is not exactly correct to say students under the AP system are not affected by the effort of other students. Since we compare these two systems under the same admission rate, the authority of the AP system has to "guess" correctly students' effort in order to set the

[19] Under our assumption that the score distributions belong to the logistic distribution, we can explicitly calculate and compare the MPWPs. Let σ^2 be the variance of the logistic distribution. The MPWP generated by the AP system, $f(\mu)$, equals $\pi \times \left(4\sqrt{3}\sigma\right)^{-1}$. It is greater than the MPWP generated by the RP system, $\int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) f(\hat{y}) d\hat{y}$, which equals $\pi \times \left(6\sqrt{3}\sigma\right)^{-1}$.

criterion at the right level to generate the right admission rate. In other words, students under the AP system are affected by other students' effort "implicitly" through the process the criterion is set. However, because we assume that the criterion is set before students invest any effort, we can still say that students do not have to worry about other students' effort under the AP system.

Our study on the RP system in this case is not completed, since the 1-for-2 case is not the only RP system that generates a 50% admission rate. To investigate the cases with more than 2 students, we consider the 2-for-4 case in detail, and then extend the result to the cases with larger number of students.

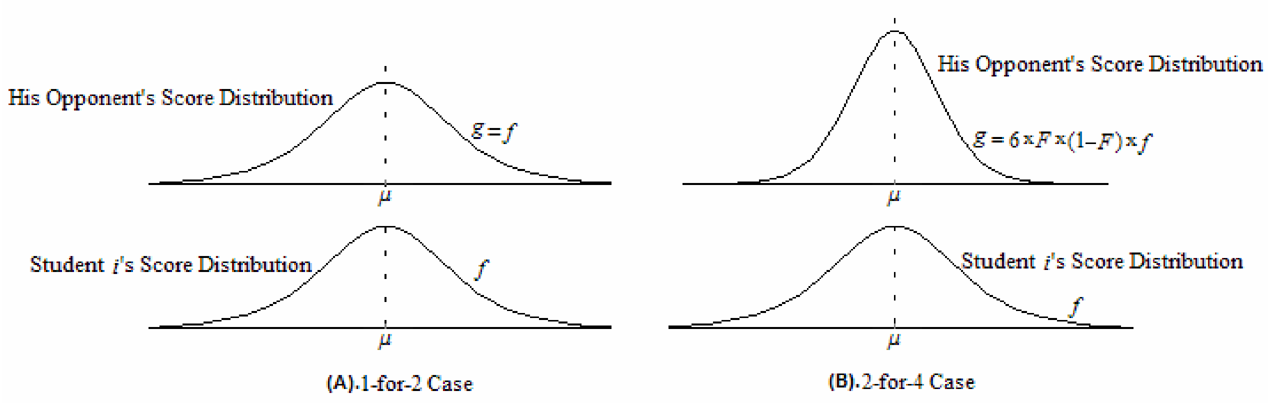
What is the difference between competing in a 1-for-2 system and in a 2-for-4 one? When a student wants to win a 1-for-2 competition, his opponent is simply the other student. When a student wants to win in a 2-for-4 competition, his imaginary opponent is the second best of the other three students. And the score distributions of the opponent in these two competitions are not the same.

We use order-statistics method to derive the score distribution of this imaginary opponent in the 2-for-4 case. The method can be introduced as follows. We randomly generate a score for each of the three opponents from their score distributions and record the value of the second best score (the middle score). We then repeat this process for many times. These recorded middle scores will form a new random variable, and this random variable is the score distribution of the imaginary opponent. In this homogeneous-student case, the process is just randomly generating three values from the same score distribution for many times, and then recording the middle of the three values drew each time to form a new random variable. In the text we will attempt to explain the result using intuition and graph, and the formal derivations are left to the appendix.[20]

We now compare the MPWP of a student i in the 1-for-2 competition to that in the 2-for-4 competition. As shown before, in the 1-for-2 case, student i's opponent's score distribution is identical to his own score distribution (the same mean and the same variance), as shown in panel (A) of < Figure 3 >, and the MPWP of his effort is $\int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) f(\hat{y}) d\hat{y}$. While in the 2-for-4 case, the opponent's score distribution would have the same mean but smaller variance, as shown in panel (B) of < Figure 3 >. The MPWP of student i's effort in the 2-for-4 case, $\int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) g(\hat{y}) d\hat{y}$, is greater than that in the 1-for-2 case. This is because student i's

[20] See Appendix D.

MPWP is the average of the pdf of his score distribution weighted by the pdf of his opponent's score distribution. The score of his opponent in the 2-for-4 case is more concentrated around the mean, assigns more weights to those large pdf values, and thus generates a greater MPWP.



< Figure 3 >

It is very straightforward to see why the score distribution of the opponent has a smaller variance in the 2-for-4 case than in the 1-for-2 case. The former distribution consists of the middle value of three random draws from a random variable, and the latter distribution comes from one single draw from the same random variable. Naturally, the middle value of three draws has a higher chance to be around the middle (and is less likely to be very large or very small) than the single draw.

We can easily apply this result to the cases with many more competitors. For example, in the 50-for-100 case, the imaginary opponent is the 50th of the other 99 students and the score distribution of this imaginary opponent consists of the middle value of 99 draws. When the number of students (N) increases, the variance of the score distribution of the imaginary opponent decreases (while the mean of the distributions stays unchanged). Thus the MPWP gets larger when the number of students gets larger: the opponent's score distribution concentrates more toward the mean and assigns more weights to the large pdf values. And when $N \to \infty$, the middle value almost always lies exactly in one value: the mean. Namely, the variance approaches zero and the score distribution of the opponent "collapses" to the mean. And the MPWP under the RP system will approach the MPWP under the AP system since all the weights are put on the pdf of the mean.

< Table 1 > shows the simulation results. We take a logistic distribution with mean 0 and variance σ^2, and we calculate the MPWPs under RP systems with different numbers of students and the MPWP under the AP system. < Table 1 > clearly confirms the results we obtained from graphic analysis.

Table 1: MPWPs under different systems with 50% admission rate

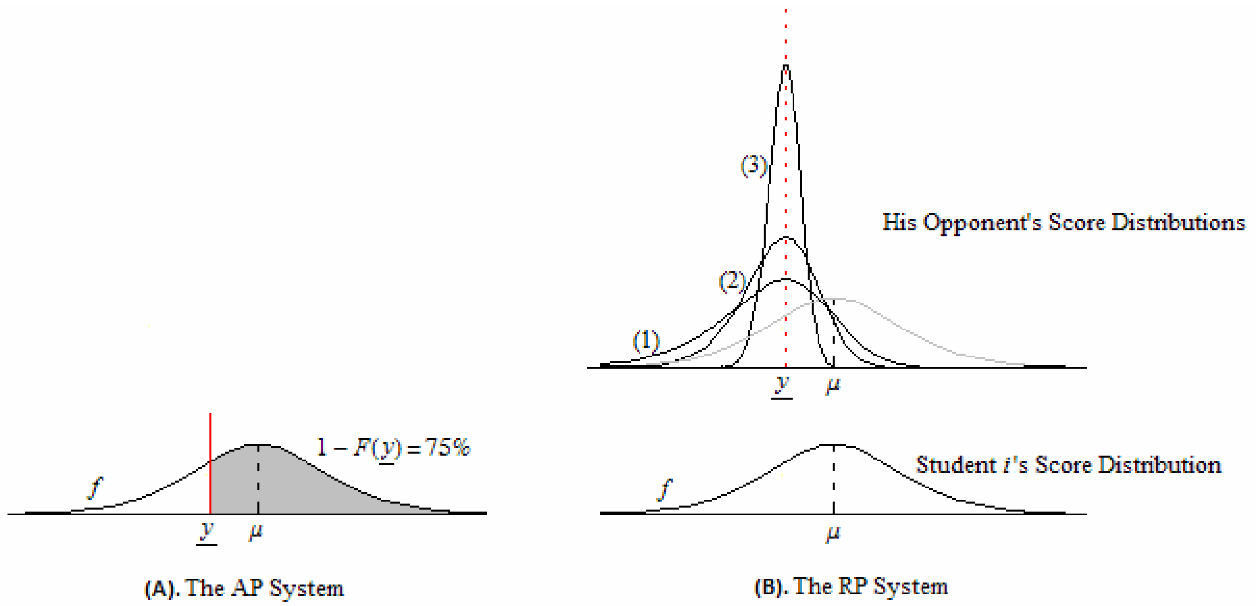
	RP ($N=2$)	RP ($N=4$)	RP ($N=10$)	…	RP ($N=100$)	…	RP ($N=1000$)	…	AP
MPWP ($1/\sigma$)	0.3023	0.36276	0.412227	…	0.44896	…	0.452997	…	0.45345

We would like to briefly summarize the results we obtained so far. In the homogeneous case with a 50% admission rate, the MPWPs generated by the RP systems are smaller than the MPWP generated by the AP system, they increase as the number of student increases, and they approach the MPWP generated by the AP system from below when the number of students approaches infinity.

2.2.2. Homogeneous students with admission rates other than 50%

Both the method and the result of the 50% case apply to the cases with admission rates other than 50%.[21] Use the 75% admission rate as an example. Under the AP system, the criterion must be set to generate an equilibrium shown in the panel (A) of < Figure 4 >: the criterion ends up at the 25 percentile of the logistic distribution whose mean equals the equilibrium level of effort. Under the RP system, the 75% admission rate means that the student in question (student i) has to beat the third best (the worst) of the other 3 in a 3-for-4 case, has to beat the 6th best of the other 7 in the 6-for-8 case, has to beat the 30th best of the other 39 in the 30-for-40 case, and so on. The score distributions of the imaginary opponents, and thus the MPWPs, can be calculated using similar method. These results are shown in Panel (B) of < Figure 4 >.

[21] We once again rely on intuitive and graphic demonstration in the text and leave formal analysis to the appendix (Appendix E).



< Figure 4 >

The bottom of Panel B shows the score distribution of student i, and the top shows the score distributions of his imaginary opponents in different cases. Distribution (1) is the score distribution of the imaginary opponent in the 3-for-4 case, distribution (2) is that of the 6-for-8 case, and distribution (3) is that of the 30-for-40 case. One can easily capture the pattern of the imaginary students' score distributions: their variances decrease as N increases and they eventually converge to a fixed value $\underline{y}$ when N approaches infinity. This pattern is similar to the 50% case, and similar pattern brings similar results. In Appendix D, we demonstrate the following three results for the 75% admission rate: the MPWP under the AP system is greater than the MPWP generated by the 3-for-4 RP system, the MPWP generated by the RP system increases as N increases, and the MPWP generated by the RP system approaches the MPWP generated by the AP system from below as $N \to \infty$.

There is actually a loose part in the link between the pattern and the result: it is not as obvious as the 50% case why the MPWP under the 3-for-4 RP system is smaller than the MPWP under the 75% AP system and why the MPWP under the 75% RP system increases as N increases. The first one is not obvious because, on the one hand, the MPWP generated by the AP system is no longer equal to the highest pdf value, and on the other hand, the MPWP under the 3-for-4 RP system actually puts some weights on the highest pdf value. And the second one is not obvious because, as the variance of opponent's score distribution decreases and the distribution concentrates toward the criterion, the weights assigned to the high pdf values decrease and this has a tendency to lower the value of MPWP. In other words, the wasted-effort explanation is less applicable when the admission rate is farther away from 50%.

However, we can prove that, under the condition that the original score distributions belong to the logistic distribution, all of the three results mentioned previously hold for the 75%-admission-rate case. We believe the force behind these results is a bias occurs in the process that students identify their imaginary opponents under the RP system. For example, in the 3-for-4 competition, the "unbiased" identification of the imaginary opponent corresponding to the 75% admission rate should be the one between the 3rd and the 4th students. In the RP system we are considering, the student in question (again we call him student i) actually takes himself out of the competition pool and identifies his imaginary opponent only from the remaining 3 students. And the opponent he identifies is the worse one of the remaining 3 students, not the "unbiased" one we just mentioned.[22] Comparing "the worst one out of 3" to "the one between the 3rd and 4th out of 4", one can see the identification is biased toward the far end. This bias brings the opponent's score distribution away from student i's score distribution, and the MPWP of student i's effort is thus biased downward. One can also see that as the number of students in the competition pool increases, the significance of this bias decreases. The values of MPWP would increase as the downward bias decreases. When $N \to \infty$, the bias becomes ignorable and the value of MPWP under the RP system approaches the MPWP value under the AP system.

In fact, we obtain a more general and powerful outcome: the three results hold not only in the 50% and 75% cases, they actually hold in all admission rates. Since comparing MPWPs gives us the outcomes of comparing effort levels, we obtain the following results.

Result 1

In the case that all students have the same ability and their score distributions belong to the logistic distributions[23],

1. the RP system induces less effort than the AP system with the same admission rate,
2. the effort induced by the RP system of the same admission rate increases as the number of students increases, and
3. the effort induced by the RP system approaches the level induced by the AP system of the same admission rate from below as the number of students approaches infinity.

Proof: (See Appendix E)

[22] See Appendix E for a graphic illustration of this point.

[23] We show in Appendix E that these results hold in the case that the measurement error belongs to the normal distribution. We cannot rule out the possibility that the RP system may induce more effort than the AP system in the homogeneous case if the assumption of the bell-shape distribution is dropped. An example that the RP system generates a higher level of MPWP is also shown in Appendix E.

Recall that in the 50% case we used "wasted effort" to explain why the RP system induces less effort; and in the 75% case we used the "identification bias" explanation. To explain the general case, **Result 1**, intuitively, we have to rely on both. When the admission rate is near 50%, the wasted-effort explanation is more relevant than the identification-bias explanation. And when the admission rate is far away from 50%, the latter becomes more pertinent. The effect of wasted effort is strong around the 50% and is weak elsewhere because the advantage of the AP system, the productivity of effort is concentrated at a certain level, is more obvious when the criterion is very close to the mean score. In contrast, the sampling bias is stronger when the admission rate is far away from 50% than when the admission rate is near 50%. From the graph shown in Appendix E, one can see that the identification bias does not even exist in the 50% case.

Besides the surprising result that the RP system induces a lower level of effort, **Result 1** also has some interesting implications regarding the exam system's effort-inducing function. First of all, we found that the effort level induced by the RP system is sensitive to how the exam is organized. More precisely, splitting a large exam pool into several smaller groups, and holding the admission rate unchanged, lowers students' effort. In the real world, college entrance exams can be and actually are organized in different ways. In Korea and Taiwan, the college entrance exam is held nationwide and students basically compete with all other students of the same academic age in the whole nation. In China, some colleges assign quotas to each province and thus students compete only with students from the same province.[24] In the U.S., many universities consider applicants' class rankings in the admission process. And those class rankings come from RP systems with much smaller numbers of students compared to those in East Asia. In Japan, some universities hold their own entrance exams at the same day and thus lower the size of the competition pool since students can only participate in one of the exams held at the same day. These different systems may be designed to serve different functions, but the way the competition pool is divided does have significant impact on students' effort. And this issue, to our knowledge, has not yet attracted people's attention.

Secondly, our result shows that, for both the AP and RP system, the highest effort is induced when the admission rate is 50%. This result implies that there is no monotonic relationship between admission rate and students' effort, and it may help explain why the recent creation of many college positions in Taiwan and Korea does not lower students' effort. We will address this issue in more detail later.

[24] Some provinces in China have larger populations than Korea and Taiwan, though.

2.3. Heterogeneous students

We now move on to the heterogeneous case. We study this case for several reasons. First, we would like to see if the result we obtained from the homogeneous case that the AP system induces more effort than the RP system sustains when heterogeneity is considered. Secondly, introducing heterogeneous ability allows us to investigate the possible problems caused by asymmetric information. And, finally, introducing heterogeneous students allows us to study another function of college entrance exam: to study how these two systems perform differently in selecting students.

To keep our analysis simple, we only explore a very special case in this paper. We assume in this special case that there are only two different types of students, that these two types split the student population evenly, and that the overall admission rate is 50%.[25] We denote these two different ability levels by A_H and A_L, and we assume that $A_H > A_L$. We sometimes use A_H and A_L to refer to the students with these types. We also denote their equilibrium levels of effort as μ_H and μ_L, their chance of winning as P_H and P_L, the pdfs of their score distributions as $f_H(\cdot)$ and $f_L(\cdot)$, and the cdfs of their score distributions as $F_H(\cdot)$ and $F_L(\cdot)$.

There are two informational structures that we would like to consider in the heterogeneous-student case. We will first consider a complete-information structure under which students know their own types and also the type of all other students in the competition pool. For simplicity we also assume that in all competition pools, students are divided evenly between the two types. We will then consider an incomplete-information case where students know only their own types and not the types of other students in the pool (but they know each of their opponents has a 50-50 chance to be either one of the two types.)

2.3.1. Heterogeneous student with complete information

We again start with the AP system. We would like to begin by sketching some properties of the solution. First of all, we expect students with different ability levels to invest different levels of effort. Secondly, because of our assumption that the overall admission rate is 50% and that students in all competition pools

[25] In the homogeneous case, there is no difference between the overall admission rate and the individual admission rate. And thus, in that case, in comparing the AP and the RP system we can control both the overall admission rate and the individual admission rate. In the heterogeneous case, the comparison of these two systems must be conducted under a less controlled scenario: we can let these two systems operate under the same overall admission rate, but doing this in general does not imply individual students with different types face the same admission rate across these two systems.

are split evenly into these two types, we expect $P_H + P_L = 1$. Namely, exactly one will win out of a pair of an A_H and an A_L. And thirdly, we expect students of different types to have the same level of MPWP.

The first property is obvious. The first order condition would never hold if they have the same effort and thus the same MPWP.[26] The second property is derived directly from our assumptions. And this implies that the criterion must be located exactly at the middle of μ_H and μ_L, namely, located at where $f_H(\cdot)$ and $f_L(\cdot)$ intersect. This further implies that $f_H(\underline{y}) = f_L(\underline{y})$. Recall that under the AP system, the MPWP of students' effort is simply $f_i(\underline{y})$. We can easily see that the MPWPs of the two types of students are the same.

Next, we investigate the solution for this case. The optimization condition for A_H is $V \cdot f_H(\underline{y}) = \mu_H / A_H$; and that for A_L is $V \cdot f_L(\underline{y}) = \mu_L / A_L$.[27] And the third equation is what we just derived: $f_H(\underline{y}) = f_L(\underline{y})$. One can show the solution for μ_H, μ_L, and $\underline{y}$ exists and is unique.[28] The relation between μ_H and μ_L can be derived easily. Since $f_H(\underline{y}) = f_L(\underline{y})$, we must have $\mu_H / A_H = \mu_L / A_L$ from the two first order conditions. This in turn implies $\mu_H > \mu_L$ since $A_H > A_L$. There is a simple explanation for this point: the same MPWP means the same marginal benefit, and the same marginal benefit means the same marginal cost (from the first order condition). The high-ability students have lower marginal cost for any effort level than the low-ability students, so they have to invest more to "achieve" the same level of marginal cost.

For the RP system, once again we begin with the 1-for-2 case. The equilibrium actually has three properties identical to those of the AP system. First of all, for the same reason, $\mu_H \neq \mu_L$. And the second property, $P_H + P_L = 1$, is derived from the same set of assumptions. Only the third property, students of different types to have the same level of MPWP, needs further explanation. Applying our previous outcome, the MPWP of A_H's effort is $\int_{\hat{y}} f_H(\hat{y}) f_L(\hat{y}) d\hat{y}$ and the MPWP of A_L's effort is $\int_{\hat{y}} f_L(\hat{y}) f_H(\hat{y}) d\hat{y}$. These two MPWPs are

---

[26] The first order condition is $V \cdot f_i(\underline{y}) = \mu_i / A_i$. Recall that students' score distributions have the same level of variance. If students of different levels of A have the same μ and thus the same value of $f(\underline{y})$, the equations for the two types could not be balanced at the same time.

[27] Recall that we assume throughout the paper that the second order conditions and the participation conditions are satisfied. See Appendix B for detailed discussions.

[28] See Appendix F.

clearly identical. And the intuitive explanation behind this is that the marginal product of student i's effort is determined by the distance between the means of his and his opponent's score distributions, not on the absolute positions of the distributions.

The solution of this case consists of two equations: the first order condition for A_H, $V \cdot \int_{\hat{y}} f_H(\hat{y}) f_L(\hat{y}) d\hat{y} = \mu_H / A_H$, and the first order condition for A_L, $V \cdot \int_{\hat{y}} f_L(\hat{y}) f_H(\hat{y}) d\hat{y} = \mu_L / A_L$. One can show that the solution exists and is unique.[29] The relation between μ_H and μ_L can be derived in similar manner. Since $\int_{\hat{y}} f_H(\hat{y}) f_L(\hat{y}) d\hat{y} = \int_{\hat{y}} f_L(\hat{y}) f_H(\hat{y}) d\hat{y}$, we must have $\mu_H / A_H = \mu_L / A_L$. This in turn means $\mu_H > \mu_L$ because $A_H > A_L$. The explanation for $\mu_H > \mu_L$ in the AP system applies to this case: the same MPWP means the same marginal benefit and in turn the same marginal cost, and it takes high-ability student more effort to reach the same level of marginal cost.

We now compare the effort induced by these two systems. We first summarize their equilibrium conditions:

For the AP system: $V \cdot f_H(\underline{y}) = \mu_H / A_H = V \cdot f_L(\underline{y}) = \mu_L / A_L$ (3)

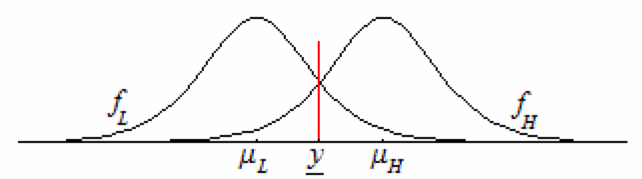
For the RP system: $V \cdot \int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_L(\hat{y}) dy = \mu_H / A_H = V \cdot \int_{\hat{y}=-\infty}^{\infty} f_L(\hat{y}) f_H(\hat{y}) dy = \mu_L / A_L$ (4)

There is one thing to notice before we conduct the comparison. From the above two equations we know that $\mu_H / A_H = \mu_L / A_L$ under both systems. This implies that ratio of the equilibrium levels of effort, μ_H / μ_L, is identical under these two systems.[30] This simplifies our analysis significantly. There are two types of students in the heterogeneous case and thus each type has to be compared individually across these two systems. However, since these two systems have the same μ_H / μ_L ratio, the comparison of one type tells us the outcome of the comparison of the other type. Therefore, it is sufficient to compare just one type. Here, we focus on comparing the effort levels of the high-ability students.

Like in the homogeneous case, the task of comparing effort levels can be simplified into one of comparing MPWPs: If a system generates a larger value of MPWP than the other system for any pair of μ_H and μ_L,

[29] See Appendix F.
[30] This, however, does not imply the levels of effort to be identical for these two systems.

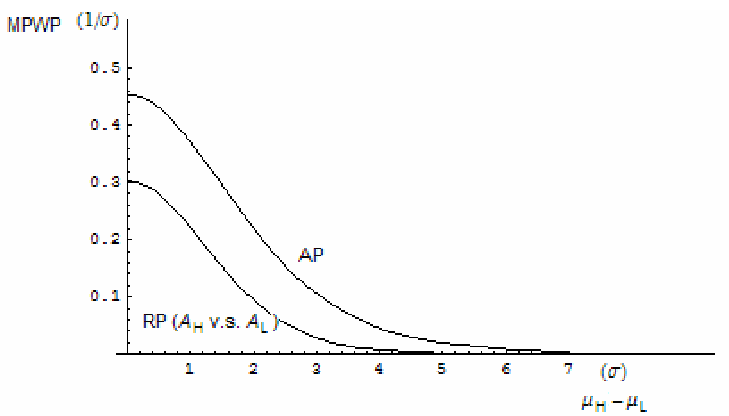
then the system would induce more effort than the other.[31] We take an arbitrary pair of μ_H and μ_L and use these two effort levels to calculate the MPWPs under these two systems. < Figure 5 > illustrates this arbitrary pair and their associated $f_H(\cdot)$ and $f_L(\cdot)$.



< Figure 5 >

The MPWP of high-ability students' effort under the AP system is $f_H(\underline{y})$, and it can be shown as the pdf of the intersection of the two score distributions in < Figure 5 >. The MPWP of high-ability students' effort under the RP system is $\int_{\hat{y}} f_H(\hat{y}) f_L(\hat{y}) d\hat{y}$: the average of $f_H(\cdot)$ weighted by $f_L(\cdot)$.

It can be shown that, under the assumption that the score distributions belong to the logistic distribution, the MPWP under the AP system is always greater than that under the RP system. < Figure 6 > shows the simulated value of MPWPs generated by these two systems. We take two logistic distributions with the same variance (σ^2) but different means (μ_H and μ_L) and calculate the MPWP under the AP system and the MPWP under the RP system for different values of effort gaps, $\mu_H - \mu_L$, measured in terms of σ. One can easily see that the MPWPs by the AP system are always greater than those by the RP system. Thus, the AP system still brings a higher level of effort than the RP system.[32]

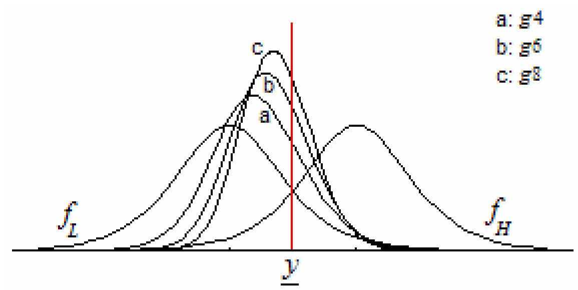


< Figure 6 >

[31] Suppose we put one system's equilibrium-effort pair into the equilibrium condition of the other and we find that the marginal benefit is greater than the marginal cost. This must mean that the second system generates a higher level of MPWP from the first system's equilibrium-effort pair. And the equilibrium for the second system can only be achieved by increasing the effort levels (and thus lowering the gap between the marginal cost and the marginal benefit).

[32] Again, we can not rule out the possibility that the RP system could generate higher effort if we consider some score distributions that have different shapes, especially those with dramatic drops somewhere in the pdf.

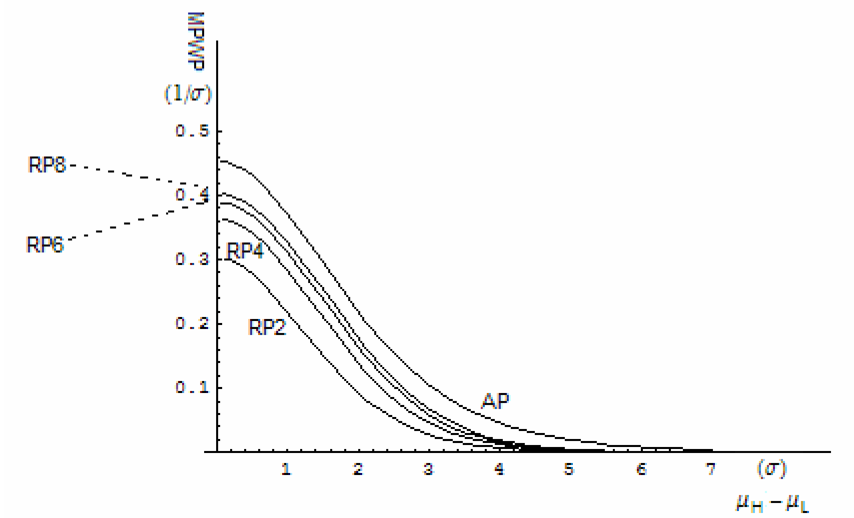
We need to consider the RP systems with a larger number of students since the 1-for-2 case is not the only RP system that brings a 50% admission rate. The result is similar to that of the homogeneous case: as N increases, the MPWP induced by the RP system increases; and when $N \to \infty$, the MPWP generated by the RP system approaches that generated by the AP system from below.



< Figure 5A >

We again leave the formal demonstration to the appendix and explain the result using intuitions and graphs here in the text.[33] Recall that the MPWP of the high-ability student's effort is the average of the pdf of his score distribution using the low-ability student's score distribution as the weights in the 1-for-2 case. When we consider a 2-for-4 case (two university positions are to be given to four students: two A_H and two A_L), the imaginary opponent of a high-ability student will be the second-best of the other three students: 2 with A_L and 1 with A_H. The score distribution of this imaginary opponent can be shown as the distribution labeled "$g4$" in < Figure 5A >.[34] The MPWP of the high-ability student's effort will be larger than that in the 1-for-2 case since the opponent's distribution is closer to his score distribution in this case and thus the weighted average gets larger. As the number of students increase from 4 to 6 and to 8, the score distribution of the imaginary opponent gets more and more concentrated toward $\underline{y}$, as shown by the distributions labeled "$g6$" and "$g8$" in < Figure 5A >. Meanwhile, the MPWPs of the high-ability student's effort will also increase, as shown in < Figure 6A >. If we increase the number of students to 1000, then a high-ability student would be competing with 499 high-ability students and 500 low-ability ones. And his imaginary opponent will be the 499th best student of the 999 opponents. What is the score distribution of this imaginary student? It is not hard for one to see that this imaginary opponent's score will almost always lie at the value $\underline{y}$, the middle value of μ_H and μ_L.[35] Namely, when $N \to \infty$, the score distribution of the imaginary opponent collapse to $\underline{y}$, and thus the MPWP approaches the pdf of his score distribution measured at that value, which is the MPWP under the AP system.

[33] See Appendix F for details.
[34] < Figure 5A > is constructed using the method similar to < Figure 5 >.
[35] See Appendix F for simulated results.



< Figure 6A >

The results can be briefly summarized as follows.

Result 2

In this special heterogeneous-students case and under the assumption that the score distributions belong to the logistic distribution, the RP system induces less effort than the AP system. The effort induced by the RP system increases as the number of students increases, and the effort level approaches the level induced by the AP system as the number of students approaches infinity. [36]

2.3.2. Heterogeneous students with incomplete information

First of all, the equilibrium of the AP system does not depend on the informational structure of students. Students do not compete with other students and thus it does not matter whether or not they know the ability of others. The equilibrium level of effort is the same as that shown in equation (3).

Like before, we start our study on the RP system with the 1-for-2 case. Consider a student with A_H. In this incomplete-information case, he only knows that there is a 50% chance he would face an A_H and the other 50% chance he would face an A_L. If he competes with an A_H, the MPWP of his effort would be

[36] These results are derived from the special case and we do not want to emphasize too much on its importance. We show these results basically to underscore the difference brought by the incomplete information. The intuitive explanation we can offer for this result is very similar to the "sampling bias" explanation: the imaginary opponents' score distributions move toward the score distribution of the student in question as N increases. And as $N \to \infty$ the opponents' score distributions collapse to the value of the criterion. Therefore, the MWPWs generated by the RP system with small N are smaller than the MPWP under the AP system.

$\int_{\hat{y}} f_H(\hat{y}) f_H(\hat{y}) d\hat{y}$; and if he competes with an A_L, the MPWP of his effort would be $\int_{\hat{y}} f_H(\hat{y}) f_L(\hat{y}) d\hat{y}$. And the MPWP in this case is simply the average value of these two MPWPs.[37]

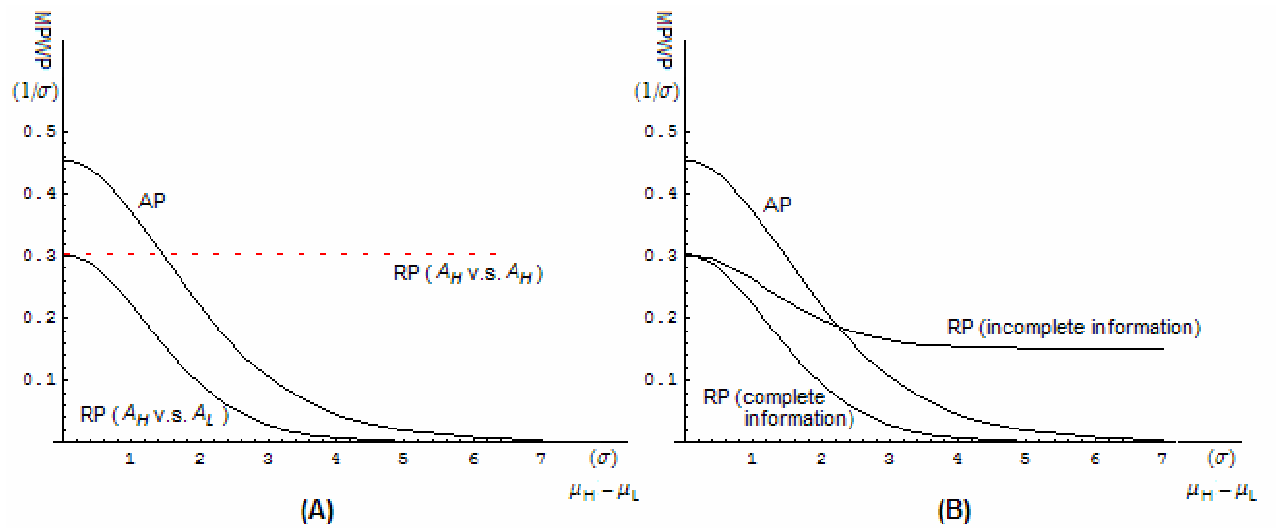
The interesting point brought by this heterogeneous-incomplete-information case is that the RP system may induce more effort than the AP system. We illustrate this point using the following figure. Panel (A) of < Figure 7 > is a duplication of < Figure 6 > with an additional dash line. Like in < Figure 6 >, the AP curve and the "RP (A_H v.s. A_L)" curve represent the MPWP of A_H's effort under the AP system and the RP system with complete information respectively. The additional line, "RP (A_H v.s. A_H)", represents the MPWP of the student's effort when he competes with another A_H. RP (A_H v.s. A_H) is obviously a straight line since the MPWP in this case is independent of the A_L's effort. From the previous paragraph, we learned that the MPWP of the effort of an A_H under the RP system with incomplete information is the average of the "RP (A_H v.s. A_L)" curve and the "RP (A_H v.s. A_H)" line, as shown in Panel (B). From Panel (B) one can see that the MPWP generated by the RP system with incomplete information is greater than the MPWP generated by the AP system when the gap between the equilibrium effort levels of these two types is large enough. This is because when the effort gap becomes substantial, both the criterion and the opponent's score distribution become farther away from the score distribution of A_H and thus the effort of A_H becomes less productive. This not only lowers the MPWPs generated by both the AP system and the RP system with complete information, but also lowers the difference between these two MPWPs.[38] And since the RP (A_H v.s. A_H) line does not decrease as the effort gap increases, sooner or later the average of the RP (A_H v.s. A_L) curve and the RP (A_H v.s. A_H) line will be greater than the AP curve. In other words, under the incomplete-information RP system, students have some chances to compete with students of their own type and that increases the expected productivity of their effort. This increase in productivity is relatively large when the effort gap is large. And when the gap is large enough this increase will reverse the order of the effort induce by the AP and the RP system.[39]

---

[37] This comes directly from our assumption that students are risk neutral.

[38] When both MPWPs get very small, the gap between them also becomes very small.

[39] This result can be obtained from another scenario. In the case of complete information, if the composition of competitors does not always have the " A_H v.s. A_L " combination but instead have other possible combinations like " A_H v.s. A_H " and " A_L v.s. A_L ", then the "expected" effort level for an A_H in such a competition would be identical to what we obtain here.



< Figure 7 >

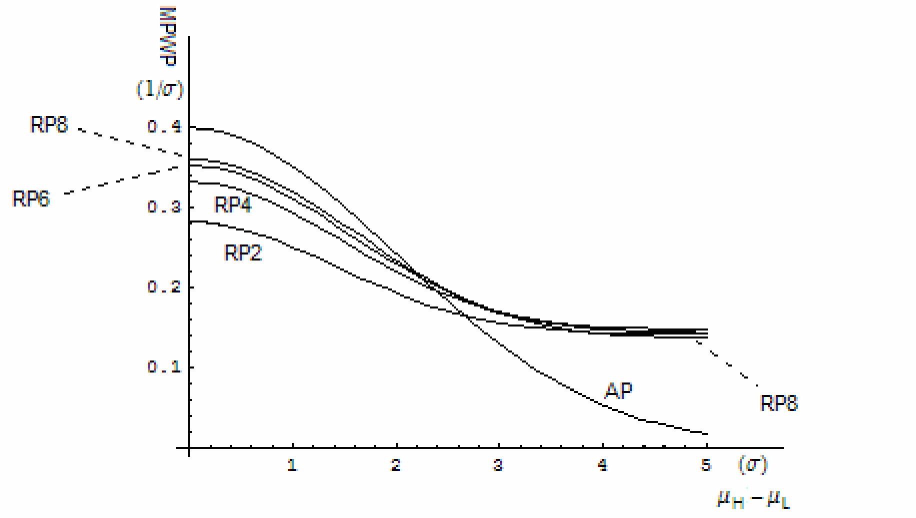
Similar method can be used to study the incomplete-information cases with $N > 2$. Take the 2-for-4 case as an example. A student with A_H will analyze his possible opponents as follows: there is a $\frac{1}{8}$ chance he will face 3 students with A_H, another $\frac{1}{8}$ chance he will face 3 students with A_L, a $\frac{3}{8}$ chance he will face 2 students with A_H and 1 with A_L, and the remaining $\frac{3}{8}$ chance he will face 1 student with A_H and 2 with A_L. The score distributions of the imaginary opponents in these situations can be calculated using the same order-statistics method, and then the MPWP of this high-ability student's effort will be the average MPWP calculated using these score distributions weighted by the associated probabilities.[40]

We show the MPWPs for the incomplete-information RP systems with more students in < Figure 8 >. The simulated results are obtained by the method similar to < Figure 6 >. The curve "RP4" in < Figure 8 > represents the MPWP for the 2-for-4 case, the "RP6" curve for the 3-for-6 case, and the "RP8" for the 4-for-8 case. We observe the following patterns. When the effort gap ($\mu_H - \mu_L$) is small, hence not very different from the homogeneous case, the MPWPs generated by the RP systems are smaller than the MPWP generated by the AP system, and they increase as N increases. And when the effort gap increases, the MPWP curves for larger N tend to drop faster than those for small N, and this reverses the order of the MPWPs. When the effort gap is large, the MPWPs generated by the RP systems are greater than the MPWP generated by the AP system, and they decreases as N increases.[41]

---

[40] See Appendix G for detailed analysis.

[41] Another pattern is about the relationship between the number of students and the required gap to make RP induce more effort. This required gap decreases as N increases for small value of N, and it increases with N when N is large.



< Figure 8 >

The result below is a brief summary of what we find from the heterogeneous incomplete-information case.

Result 3

In the heterogeneous case with incomplete information, the RP system generates more effort than the AP system when the gap between ability levels is large enough.[42] The feature of incomplete information also changes the relationship between N and the effort levels induced by the RP system. In the case of small ability gap, the RP system with larger N induces more effort. When the ability gap becomes large enough, the RP system with larger N induces less effort.

2.4. Summary and discussions for section 2

Under the assumption that the random factor of students' scores belongs to the logistic distribution and the assumptions that we made to construct the special heterogeneous case,[43] we obtain the following results:

1. In the homogeneous case, the AP system generates more effort than the RP system. The effort level induced by the RP system increases with N, and approaches the level generated by the AP system from below as $N \to \infty$.
2. In the heterogeneous case with complete information, the same results as those in (1) are obtained.
3. In the heterogeneous case with incomplete information, the RP system generates more effort than the AP system when the difference between students' ability levels is large enough.

[42] A larger ability gap generates a larger effort gap. From equation (3) and (4) in page 23, we can obtain $(A_H - A_L) = (\mu_H - \mu_L)/(V \cdot MPWP)$. When $(A_H - A_L)$ becomes larger, $(\mu_H - \mu_L)$ must also be larger: If $(\mu_H - \mu_L)$ gets smaller, then the MPWP would get larger. The RHS becomes smaller and the equation will not hold.

[43] We would like to remind the reader that the assumptions behind this special case are that students are split evenly into two types and that the overall admission rate is 50%.

4. In the heterogeneous case with incomplete information, the effort generated by the RP system increases with N when the ability difference is small; and the effort generated by the RP system decreases when N increases when the ability difference is large.

Intuitively, the reason behind the first two results is that the marginal product of the effort spent in the attempt to beat other students depends on the opponent's realized score. Some of those realized scores bring very low productivity to student's effort. Those low-productivity scores bring down the averaged marginal productivity and thus lower the level effort invested. When N is small, the chance for these low-productivity scores to happen is relatively large. And this chance decreases as N increases. When N is extremely large, the opponent's score distribution is almost equal to the criterion that has to be set by the corresponding AP system and these two systems induce the same level of effort. The third result is derived directly from two outcomes of our model. First, the productivity of a student's effort is larger when he faces students with the same ability than with different ability. Secondly, the actual combination of students in the competition pool only affects students' effort under the RP system, not the AP system. And of course these two factors become more significant when the informational problem gets stronger due to a larger ability difference. The fourth result comes from the power-wrestling of two effects caused by the increase in N. On the one hand, like what works in the first two results, an increase in N lowers the chance that effort being wasted on low-productivity scores and thus tends to increase the effort caused by the RP system. On the other hand, an increase in N lowers the informational problem in the heterogeneous case caused by uncertainty over the composition of the competition pool and thus tends to lower the effort level.[44] The second effect is small when ability gap, and thus the informational problem, is small, so the first effect dominates. When the ability gap and the informational problem get larger, the second effect starts to dominate at some point.

These results have some interesting implications.

1. The norm-referenced feature of the exam (the RP system) alone should not be blamed for the problem that students study "too hard". Trying to beat other students induces more effort than trying to beat a criterion only when students are asked to compete without knowing other students' abilities.

---

[44] For example, in the 1-for-2 case the uncertainty is severe: the student either faces an A_H or an A_L, and the marginal product of effort is very different from competing in these two cases. In the, say, 100-for-200 case, the student knows that the chance he would be competing with a combination around the neighborhood of "199 A_H and no A_L" combination is very small. In the latter case, the variance of the marginal productivity of students' effort is smaller. It is in this sense we said that the information problem is smaller when N gets larger.

2. The effort-inducing function of the RP system is sensitive to both the changes in the size of competition pool and to the informational structure. The effort-inducing function of the AP system is independent of these changes. Students' effort on preparing for the college entrance exam can be lowered effectively by dividing the nationwide exam into several local ones (but holding unchanged the admission rate and the distribution of ability) and at the same time offering students better information about their opponents. As we discussed before, several methods can be and are practically used to divide the competition pool. To reduce the informational problem is probably a more complicated and difficult task. Some of the methods we can think of are related to reducing the size of the competition pool. If the pools are downsized to a high school or even a class, then it would not be too hard for students to figure out their opponents' abilities from studying together for years. If the pool is divided into larger groups such as counties or universities-to-attend, then some test institutions might be able to create and provide this kind of information to students. If the exam is still nationwide, letting students take some mock exams and offering detailed analysis of the test results could help students know their opponents better. Grouping students according to their ability into different classes or even different schools might also help when the exam is nationwide.
3. Changing the size of the competition pool and/or the informational structure of a RP system can make the system generate different levels of effort. In some occasions, the RP system can be manipulated to generate a range of effort that includes the level generated by the AP system. This implies that, in these occasions, we can claim that the RP system is superior to the AP system without knowing the authority's preference: no matter what the most desirable effort level is, the RP system can be manipulated to generate an effort level close or even equal to the most desired level.

3. Aptitude Test versus Achievement Test

In this section we explore the other feature of exam that may have impact on students' effort: the achievement-test feature. Notice that what we have considered so far is a test that the ability level does not directly affect the score distribution. To introduce the aptitude-test feature into our model, we change the representative student's (student i's) score function from $y_i = \mu_i + \varepsilon_i$ to $z_i = \alpha \cdot A_i + y_i = \alpha \cdot A_i + \mu_i + \varepsilon_i$. In the new function, μ_i is still student i's effort, ε_i is still the measurement error associated with the exam. The new term is αA_i, which is to capture the direct impact of student i's ability on his score. One can think adding this additional term as asking some additional IQ-type questions that students with high ability can answer correctly without spending any time preparing for them and low ability students cannot answer no matter how much time they spend on studying for them.

We assume $\alpha \geq 0$. When α is small, students' (mean) scores are mainly determined by their effort; and when α is large, students' ability plays a more dominant role in their scores. That is, our previous score function, an "pure" achievement case,[45] can be considered as a special case of this general form with $\alpha = 0$. And as α gets larger, the measure of students' learning achievement is more "biased" by their ability. We call the exams with positive values of α the "partial" achievement tests.

We now study how this change in the score function affects our results. To avoid confusion, we call the pdf and cdf of student i 's score distribution $h_i(\cdot)$ and $H_i(\cdot)$ respectively in the partial-achievement case.[46]

First of all, shifting the exam from the pure-achievement test to the partial-achievement test does not affect the result we obtained from the homogeneous case. For the AP system with a 50% admission rate, the equilibrium conditions are

$$\begin{cases} V \cdot h_i(\underline{z}) = \mu_i / A \\ \underline{z} = \alpha A_i + \mu_i \end{cases} \quad \text{,or equivalently } V \cdot h_i(\alpha A_i + \mu_i) = \mu_i / A .$$

And the comparable condition under the pure-achievement test is

$$V \cdot f_i(\mu_i) = \mu_i / A .$$

One can easily see that the equilibrium levels of μ_i derived from these two conditions are identical: since $\alpha A_i + \mu_i$ is the mean of random variable represented by $h_i(\cdot)$ and μ_i is the mean of the random variable represented by $f_i(\cdot)$, both $h_i(\alpha A_i + \mu_i)$ and $f_i(\mu_i)$ equal the value of pdf of the score distribution measured at the mean. By intuition, adding ability to the score is like giving everyone an identical amount of free points. And to hold the admission rate stay unchanged, the criterion has to be raised by the same amount. Since everyone receives an identical amount of free points and sees the criterion increased by the same amount, nothing changes. The same applies to all other admission rates under the AP system:

[45] It is not very precise to call this kind of test a "pure" achievement test. In this kind of test, the ability levels affect students' scores indirectly: through affecting the cost of studying. Also, even if an exam is designed to be a pure-achievement one, it is very hard to prevent students with high ability from getting advantage in the exam. High-ability students tend to read more rapidly, understand questions better, analyze and calculate faster, and present answers better. If exams have too many questions and/or have some questions that are not very closely confined to the curriculum and thus require general analyzing skills, a substantial part of the score would be determined by students' ability. The term "pure" here can be more precisely interpreted as the test which does not contain any IQ-type questions and high-ability students have no advantage other than having lower studying cost. Namely, in this "pure" achievement test, all questions can be prepared through studying and students investing the same effort get the same mean score.

[46] The second order conditions and the participation conditions are again assumed to hold, and thus our analyses focus on the first order conditions. The solutions in this section have the same structure as the solutions in the previous section, and thus the existence and uniqueness of the solutions can be demonstrated using the same approach.

increasing everyone's score and criterion by the same amount does not have any impact on students' behavior.

For the 1-for-2 RP system, the equilibrium conditions are

$$\begin{cases} V \cdot \int\limits_{\hat{z}=-\infty}^{\infty} h_1(\hat{z}) h_2(\hat{z}) d\hat{z} = \mu_1 / A \\ V \cdot \int\limits_{\hat{z}=-\infty}^{\infty} h_2(\hat{z}) h_1(\hat{z}) d\hat{z} = \mu_2 / A \end{cases},$$

or simply $V \cdot \int\limits_{\hat{z}=-\infty}^{\infty} h(\hat{z}) h(\hat{z}) d\hat{z} = \mu / A$ if we use the symmetric condition $\mu \equiv \mu_1 = \mu_2$ and $h(\cdot) \equiv h_1(\cdot) = h_2(\cdot)$ to simplify the equations. Compare this condition to that in the pure achievement test

$$V \cdot \int\limits_{\hat{y}=-\infty}^{\infty} f(\hat{y}) f(\hat{y}) d\hat{y} = \mu / A,$$

we can again see that the equilibrium levels of μ_i derived from these two conditions are identical: since $h(\cdot)$ and $f(\cdot)$ are the same except that they have different means, the value of $\int\limits_{\hat{z}=-\infty}^{\infty} h(\hat{z}) h(\hat{z}) d\hat{z}$ must equal the value of $\int\limits_{\hat{y}=-\infty}^{\infty} h(\hat{y}) h(\hat{y}) d\hat{y}$. The intuitive explanation here is similar to that for the AP system: giving every student the same amount of free points does not change the nature of the competition.

One can easily extend this result to the other RP systems with 50% admission rate and all other RP systems with different admission rates. In the homogeneous case, adding the ability factor into the score is nothing but giving everyone the same free points, and doing so has no impact on students' behavior. Or, to put this differently, the MPWP in the AP system depends on the relative position between students' mean score and the criterion, and the MPWP in the RP system depends on the relative position of students' mean scores. Adding ability to score does not change those relative positions, and thus the equilibrium effort levels derived from the pure-achievement test can automatically satisfy the equilibrium condition for the partial-achievement test.

Adding ability directly into the score function, however, does change the equilibrium effort level in the heterogeneous case. The bottom line is that students of different abilities now get different amounts of free points.

First, let's consider the AP system. For the special case with heterogeneous students under complete information, the equilibrium conditions are

$$\begin{cases} V \cdot h_H\left(\underline{z}\right) = \mu_H / A_H \\ V \cdot h_L\left(\underline{z}\right) = \mu_L / A_L \end{cases}$$

Applying the 50%-admission-rate condition $h_H\left(\underline{z}\right) = h_L\left(\underline{z}\right)$, we get $V \cdot h_H\left(\underline{z}\right) = \mu_H / A_H = V \cdot h_L\left(\underline{z}\right) = \mu_L / A_L$.

And $V \cdot f_H\left(\underline{y}\right) = \mu_H / A_H = V \cdot f_L\left(\underline{y}\right) = \mu_L / A_L$ is the corresponding condition for the pure-achievement test.

Since we need to consider two effort levels in each case, to compare effort levels in the heterogeneous case again is more complicated. However, similar to our previous analysis, we can see the μ_H / μ_L ratios in these two cases equal to the same value (A_H / A_L), and thus we can simply compare the values of μ_H across these two cases and the same relationship holds for the values of μ_L. We compare the effort level in these two cases by the following steps: first substitute the solution of the pure-achievement test into the condition for the partial-achievement test and check whether the condition is satisfied. If the condition is not satisfied, we would then check whether the effort level needs to be increased or decreased to satisfy the condition.

Let's denote the equilibrium effort levels from the pure-achievement case as μ_H^* and μ_L^*. The MPWP in that case, $f_H\left(\underline{y}\right) = f_L\left(\underline{y}\right)$, depends on the gap between them. When we substitute μ_H^* and μ_L^* into the first order condition for the partial-achievement case, the value of MPWP will be smaller. This is because the MPWP in the partial case is determined by the gap between $\alpha A_H + \mu_H$ and $\alpha A_L + \mu_L$, not just μ_H and μ_L.[47] μ_H^* and μ_L^* are thus "too high" to sustain in the partial-achievement case. To maintain the equilibrium, the effort levels must be lowered. Thus, introducing ability into the score function lowers the equilibrium level of effort.

We can also investigate the impact of introducing ability into the score function on the selection function of the exam system. We compare the selection functions of different exam systems through investigating which

[47] To be more precise, the MPWP for A_H is determined by the gap between $\alpha A_H + \mu_H$ and the criterion $\underline{z}$. But since $\underline{z}$ must locate in the middle of $\alpha A_H + \mu_H$ and $\alpha A_L + \mu_L$ in the case we are considering, we can also say that the MPWP is determined by the gap between the two mean scores.

system gives the high-ability students a higher chance to win a college position. Namely, whichever system has a higher level of P_H is considered as selecting better.[48] The selection function depends entirely on the gap between the means of the score distributions. Even though the levels of the effort falls in this case and thus the effort gap decreases, the newly introduced ability gap in the score function actually dominates and raises the gap in mean scores. This point can be illustrated using the method similar to the previous analysis but from the reverse direction. Take the mean gap from the pure-achievement case (which is nothing but the effort gap) and apply it to the partial-achievement case. These two cases will have the same MPWP and thus must also have the same effort level. This is not consistent with the initial condition that these two equilibria have the same gap in mean scores: the same equilibrium levels of effort bring a larger gap between mean scores in the partial-achievement case than in the pure-achievement case. This is again because the additional ability gap in the partial-achievement case. In other words, the MPWP generated by the pure-achievement case's mean gap is "too large" for the partial-achievement case. Thus, the MPWP in the partial-achievement case must be smaller than the MPWP in the pure-achievement case. This in turn means the gap between mean scores must be larger, and so the selection function becomes better.

The analyses and results of the AP system apply well to the 1-for-2 RP system. The equilibrium conditions for the 1-for-2 RP system are

$$\begin{cases} V \cdot \int\limits_{\hat{z}=-\infty}^{\infty} h_H(\hat{z}) h_L(\hat{z}) d\hat{z} = \mu_H / A_H \\ V \cdot \int\limits_{\hat{z}=-\infty}^{\infty} h_L(\hat{z}) h_H(\hat{z}) d\hat{z} = \mu_L / A_L \end{cases}$$

And the corresponding conditions for the pure-achievement case are

$$\begin{cases} V \cdot \int\limits_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_L(\hat{y}) d\hat{y} = \mu_H / A_H \\ V \cdot \int\limits_{\hat{y}=-\infty}^{\infty} f_L(\hat{y}) f_H(\hat{y}) d\hat{y} = \mu_L / A_L \end{cases}$$

The same tricks can be used to compare the effort levels. Suppose we substitute the effort levels from the pure case, μ_H^* and μ_L^*, into the conditions of the partial case, then the MPWPs in the partial case would be too small to sustain the condition. And again this is because the additional ability gap in the partial case increases the mean-score gap and bigger gap generates lower MPWP. One can also show that the partial case selects better using similar method.

48 See section 4.1 for a more detailed discussion of the selection function.

It is not hard to see that this result applied to the RP cases with different N and also to the RP systems under incomplete information. The same logic applies: the effort level of the pure-achievement case is too high for the partial-achievement case, and the gap between mean scores in the pure case is too small for the partial case. Thus the partial-achievement case has a smaller gap in effort levels and a larger gap in mean scores in equilibrium.

Result 4

Introducing ability into the score function has no impact on the effort-inducing function of the exam in the homogeneous case. In the heterogeneous case, doing so lowers students' effort and improves the selection function of the exams.[49]

Regarding the difference in the effort-inducing functions between the AP system and the RP system, one can easily see that all the results we obtained from the pure-achievement case still holds in the partial-achievement case. In the homogeneous case, this point is obvious since adding aptitude-factor does not change the effort levels. In the heterogeneous case, since the equilibrium conditions for both the AP system and the RP system in the partial-achievement case have the same form with the conditions in the pure-achievement case, the method we used to compare the effort-inducing functions between the AP and the RP systems can be applied here. And the same results will be obtained.

4. Applications

4.1. The selection function of the exam systems

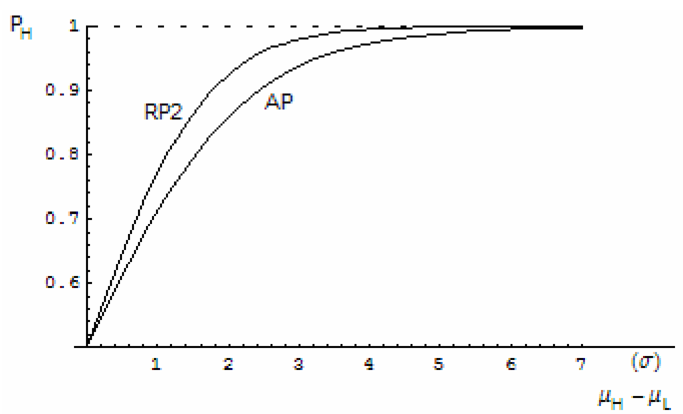
In this application we compare the selection function of the AP system and the RP systems. First of all, we need to define a measure to evaluate and compare the performances regarding the selection function. The most straightforward measure is $(P_H - P_L)$, the gap between the winning probability of A_H and that of A_L.[50] And, in this two-type case we are considering, a larger P_H implies a smaller P_L and in turn implies a larger $(P_H - P_L)$. In other words, we say a system performs the selection function better if under the system the high ability students' winning probability is greater than that under the other system. To focus on the difference caused by these two systems, we return to the pure-achievement case. Also, we again concentrate

[49] From this result, we can also claim that lowering the weights of the aptitude-test factor in the exam raises students' effort. This is consistent with the claim of some people in the U.S. that lowering the importance of the SAT score in the admission process can stimulate more effort from high school students.

[50] We would need a more complicated measure such as stochastic dominance if we have more than two types.

on the special case: the case that there are two types of students, that they split the population evenly, and that the overall admission rate is 50%.

We first consider the AP system and the 1-for-2 RP system. How to compare the winning probabilities generated by these two systems? Under either system, the probability gap obviously depend on the effort gap (the gap between μ_H and μ_L). What needs to be explored further is that whether the two systems generate the same probability gap from the same effort gap. < Figure 9 > shows the simulation result. We take two score distributions (two logistic distributions with the same variance but different means) and calculate the winning probabilities of the high-effort student for different effort gap ($\mu_H - \mu_L$) measured in terms of the standard deviation of the distribution under these two systems. The pattern can be seen clearly from the figure: the RP system in general generates a larger probability gap than the AP system. The difference between these two systems is close to zero when the effort gap is either very small (P_H and P_L are both very close to 0.5 under both systems) or very big (P_H is close to 1 and P_L close to 0 under both systems). When the effort gap is of a moderate size, the RP system selects substantially better than the AP system given the same effort gap.



< Figure 9 >

One can easily see that this probability gap becomes smaller when N (the number of students) increases in the RP system.[51] This is because the score distributions of the imaginary opponent in the RP system converge to the criterion of the AP system as N increases.[52] For the incomplete-information cases, the only comparable case is the one that the two types actually split the population evenly (but students do not know this). And in this case, all of the results we obtained from comparing the AP and the complete-information RP system apply.

---

[51] See Appendix H for a graphic illustration.
[52] See < Figure 5A > in page 25.

The comparison of the selection function of these two systems becomes complicated when we consider both the effort-gap aspect and the probability-gap aspect. One the one hand, the AP system generates a larger effort gap than the RP system in the complete information case and in the incomplete information case with small ability gap. On the other hand, the RP system generates a larger probability gap from the same effort gap when N is small and students are moderately different in ability. We make an attempt to combine these two aspects and obtain the following outcomes. First, when N is small and students are not very different in ability, the AP system should select better. Secondly, the RP system may select better when N is small, students are fairly different in ability, and students do not know the ability of their opponents. Thirdly, when N is large, the probability gap becomes ignorable and the selection function is determined by the effort gap. However, the precise comparisons can only be conducted when such factors as the number of students, their ability gap, and the informational structure are clearly given.

4.2. Will increasing admission rate mitigate the suffering in the exam hell?

The other interesting application of our model is to study whether or not an increase in the number of university positions eases the suffering under the exam hell. Both Taiwan and Korea expanded their tertiary education capacity and increased college admission rate dramatically in recent years. Many think the increase in supply should have eased the competition and thus lessened the sufferings in the exam hell. However, there is no evidence that this increase in admission rate lowers students' effort in these countries. Our model offers a comprehensible explanation for the relationship between students' effort and the admission rate, and it can help us understand the seemingly puzzling phenomenon we just described.

According to our model, under the RP system, the MPWP of a student's effort is determined by how close his score distribution is to his opponent's score distribution: the closer the two distributions to each other, the larger the MPWP.[53] The effect of some additional college positions (or, equally, the effect of increasing admission rate) on students' effort can be derived easily from this insight. If the additional positions bring closer the imaginary student's score distribution, they would make students work harder.

In the homogeneous case of our model, whether or not an additional position would bring the opponent's score distribution closer depends on whether the additional position moves the admission ratio toward or away from 50%. It can be easily shown that the MPWP for a give number of students is maximized when

[53] "Close" is not a very precise term here. Actually both the means and variances of his score distribution and his imaginary opponent's score distribution are relevant. To be more precise, we should say that the more weights the opponent's score distribution gives to the high-pdf-value parts of his own score distribution, the larger the MPWP.

the admission rate is 50%. The same idea applies to the heterogeneous case with complete and incomplete information.[54]

The importance of the 50% should not be over-emphasized. The 50% is critical in our model regarding this issue because we assume that all college positions are identical to every student and that (in the heterogeneous case) the student population is split evenly between high and low types. Allowing college positions to be different and/or considering different composition of the student population will alter the admission rate that maximizes effort. Therefore, there is no general “critical value” of admission rate that can be derived without studying the structure of students’ ability and college positions.

In short, expanding university capacity could either raise or lower students’ effort depending on whether it makes students’ effort more or less productive. And one needs to study carefully the composition of students and their preferences over college positions in order to know toward which direction it moves.

5. Concluding Remarks

In this paper we investigate the relationship between the over-studying problem associated with college entrance exams in East Asia and their designs. We found that the norm-referenced feature of the exams does not contribute much to the over-studying problem as many expected. Compared to the alternative, the criterion-referenced test, the norm-referenced one in general induces less effort except when students are very different in their ability and the abilities of the opponent students in the competition pool is not known.

We find that the effort-inducing function of the norm-referenced exam is sensitive to both the size and the informational structure of the competition pool. This result has two implications. First, this flexibility gives the norm-referenced exam (a rewarding mechanism based on relative performance) an edge over the criterion-referenced exam (a rewarding mechanism based on absolute performance). Second, this implies that the exam authority using the norm-referenced system can effectively lower students’ effort by lowering the sizes of the competition pools and at the same time improving the informational structure of the competition pools.

We find that the other feature of the college-entrance exams in East Asia, the achievement-test feature, may be responsible for the over-studying problem that concerns many in those countries. Thus the other possible

[54] See Appendix I for detailed analysis.

way for the educational authority to lower students' effort is to increase the portion of the score that is determined by students' ability and is independent of students' effort.

We apply our method to compare the selection function of the criterion-referenced and norm-referenced tests, and we find the criterion-referenced test selects better when the number of students competing in the group is small and students are similar in their abilities. And the RP system may select better when the number of students is small, students are fairly different in ability, and students do not know the ability of their opponents. We also study the relationship between admission rate and students' effort and find that the relationship is not monotonic. Namely, our theory offers an explanation for the phenomenon happened in East Asia that the creation of more college positions does not release stress from students.

In short, we find that neither shifting to the criterion-referenced system nor adding more college positions can unambiguously lower students' effort. The approaches that we think are more effective in releasing students' suffering in the exam hell are increasing the aptitude-determined portion of the score, reducing the informational problem about the composition of the competition pools, and at the same time lowering the size of the pools.

We regard our research as an extension of the work by Becker and Rosen (1992). For us, their research consists mainly two parts: a study on the criterion-referenced and norm-referenced exams and a study on the political economy of the criterion-referenced system. In this paper we push the first part of their work further: we study and compare the effort-inducing function and the selection function of the criterion- and norm-referenced tests in more details. We also plan to extend the second part of their research: to study the political economy of the exam systems. This paper can be expanded naturally to study the impacts of the change in exam systems and/or changes in the detailed designs of the exam systems on the welfare of students with different abilities. By charactering the population's preference over those welfare changes and dealing with different political systems, we may be able to start predicting and explaining the designs and the changes in designs of the college-entrance-exam systems.

Reference

Ando, M. (2004) "Division of a Contest with Identical Prizes", *Journal of the Japanese and International Economies*, Vol. 18, 282-297.

Becker, W. and S. Rosen, (1992), "The Learning Effect of Assessment and Evaluation in High School", *Economics of Education Review*, Vol. 11, No. 2, 107-118.

Berwick, R. and S. Ross (1989), "Motivation after Matriculation: Are Japanese Learners of English still Alive after Exam Hell?" *JALT Journal*, Vol. 11, 193-201.

Bishop, J. (1998), "The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement", *Journal of Economic Education*, Vol. 29, No. 2, 171-182.

Costrell, R. (1994), "A Simple Model of Educational Standards", *American Economic Review*, Vol. 84, No. 4, 956-971.

David, H. (1984), *Order Statistics*, Wiley.

Dye, R. (1984), "The Trouble with Tournaments", *Economic Inquiry*, Vol. 22, 147-149.

Ellinger T. and G. Beckham, (1997), "South Korea: Placing Education on Top of the Family Agenda", *Phi Delta Kappan*, Vol. 78, No. 8, 624-25.

Feuer, M. and K. Fulton, (1994), "Educational Testing Abroad and Lessons for the United States", *Educational Measurement: Issues and Practice*, Vol. 13, No. 2, 31-39.

Gibbons, R. (1992), *Game Theory for Applied Economists*. Princeton University Press.

Green, J. and N. Stokey, (1983), "A Comparison of Tournaments and Contracts", *Journal of Political Economy*, Vol. 91, 349-364.

Harbaugh, R. (2001), "Achievement vs. Aptitude: the Incentive-Screening Tradeoff in College Admissions", Working Paper.

Holman R. (1991), "Exam Hell Linked to Depression", *Wall Street Journal*, 26 December, p.A4.

Knoeber, C. (1989), "A Real Game of Chicken: Contracts, Tournaments, and the Production of Broilers", *Journal of Law, Economics, and Organization*, Vol. 5, No. 2, 271-292.

Kubiszyn, T. and G. Borich, (2003), *Educational Testing and Measurement: Classroom Application and Practice*, Seventh Education, John Wiley & Sons, Inc.

Lazear, E. and S. Rosen, (1981), "Rank Order Tournaments as Optimal Labor Contracts", *Journal of Political Economy*, Vol. 89, 841-864.

Larson R. and S. Verma, (1999), "How Children and Adolescents Spend Time across the World: Work, Play and Development Opportunities", *Psychological Bulletin*, Vol. 125, No. 6, 701-736

Lee, M. (2003), "Korean Adolescents' 'Examination Hell' and Their Use of Free Time", *New Directions for Child and Adolescent Development*, No. 99, 9-21.

Lee, M. and R. Larson, (1999), "The Korea 'Examination Hell': Long Hours of Studying, Distress, and Depression", *Journal of Youth and Adolescence*, Vol. 29, No. 2, 249-271.

Madaus, G. (1991) "The Effects of Important Tests on Students: Implications for a National Examination System", *Phi Delta Kappan*, Vol. 73, No. 3, 226-231.

Nalebuff, B. and J. Stiglitz, (1983), "Prizes and Incentives: Towards a General Theory of Compensation and Competition", *Bell Journal of Economics*, Vol. 14, 21-43.

Noah, N. and M. Eckstein, (1989a), "Tradeoffs in Examination Policies: An International Comparative Perspective", *Oxford Review of Education*, Vol. 15, No. 1, 17-27.

Noah, N. and M. Eckstein, (1989b), "Forms and Functions of Secondary-School-Leaving Examinations", *Comparative Education Review*, Vol. 33, No. 3, 295-316.

Noah, N. and M. Eckstein, (1993), *Secondary School Examinations: International Perspectives on Policies and Practice*, Yale University Press.

Nunnally, J. (1964), *Educational Measurement and Evaluation*, MaGraw-Hill, Inc.

Ogura Y. (1987), "Examination Hell: Japanese Education's Most Serious Problem", *College Board Review*, 144, 8-11, 26-29.

O'Keeffe, M., W. Viscusi, and R. Zeckhauser, (1984), "Economic Contests: Comparative Reward Schemes", *Journal of Labor Economics*, Vol. 2, no. 1, 27-56.

Ono, H. (2006), "Does Examination Hell Pay off? A Cost-benefit Analysis of 'Ronin' and College Education in Japan." *Economics of Education Review*, forthcoming.

PISA (Program for International Student Assessment), (2003), *Learning for Tomorrow's World: First Results from PISA 2003*.

Sax, G. (1989), *Principles of Educational and Psychological Measurement and Evaluation*, Third Edition, Wadsworth Publishing Company.

Stevenson, H. and S. Lee, (1997), "International Comparisons of Entrance and Exit Examinations: Japan, United Kingdom, France, and Germany", Report to the U.S. Department of Education, Office of Educational Research and Improvement.

Stiglitz J. (1975), "The Theory of 'Screening,' education, and the Distribution of Income", *American Economic Review*, Vol. 65, No. 3, 283-300.

Appendix

A. Korean high-school students' long studying hours

I. Larson and Verna (1999) studied how adolescent spend their time. Table 3 of their paper summarized the results of several researches on adolescents' "total time in school work" from several countries. The data of all those researches are obtained from the experience sampling method (ESM). According to different researches on "high school" students, Korean students spent 5.6 hours per day, Japanese students spent 5.4 hours, Italian students spent 4.6 to 4.8 hours, and American students spent 3.7 hours. (The research on Korean students is limited to a time span from 9 AM to 9PM, and that makes the reported number underestimate the real schoolwork hours.) From a research concentrated on "12th grade" students, Korean students spent 7.8 hours per day, and the American students spent 2.9 hours.

II. Table 5.14 in PISA (2003) reported "student learning time" across OECD and its partner countries. The term "student learning time" is defined as "students' reports of the average number of hours spent on the "out-of and in-school" activities during each school week." The number for Korea is about 50 hours, and the average number of the OECD countries is around 35 hours. (To report a short list of selected countries: around 32 hours for Germany, 34 for France, 41 for Italy, and 33 for the U.S.)

B. The assumptions for the participation condition and the second order condition

B1. The participation condition

This participation condition is a big and seemingly real issue in the rent seeking literature. For example, consider a rent-seeker (player 1) participates in a two-player contest that gives him a profit function: $V \times \frac{x_1^3}{x_1^3 + x_2^3} - x_1$. Where V is the value of the rent, x_1 is player 1's "bid" for the rent, and x_2 is the bid of the opponent (player 2). The term $\frac{x_1^3}{x_1^3 + x_2^3}$ represents player 1's chance of winning and thus is often referred as the contest success function (CSF). The Nash equilibrium in this contest is $x_1^* = x_2^* = x^* = {}^{3V}\!/_{4}$. This equilibrium, however, does not satisfy the participation condition since both players has to spend ${}^{3}\!/_{4} \cdot V$ to get a half chance of winning V. In this case, there is no pure-strategy Nash equilibrium.

This sort of problem is not as serious in our model. The reason is that our model does not link the equilibrium level of investment to the value of the prize as closely as the rent-seeking model. The rent-seeking model ties them closely because of the following two properties. Firstly, the first derivative of the CSF (with respect to x) in this kind of form always shows up in the Nash equilibrium as the reciprocal of x multiplied by a constant. Secondly, the linear form of the bidding cost gives a constant marginal cost. Combining these two properties, the Nash equilibrium always looks like $V \cdot {}^{\alpha}\!/_{x} - \beta = 0$ (or $x = {}^{\alpha}\!/_{\beta} \cdot V$) ,where both α and β are a constant. The reason that the constant-marginal-cost is prevailing in the rent-seeking model is that many of those researches are lobby-like contests where money invested is supposed to have constant marginal product.

The first order derivative of the CSF function in our model, the MPWP, is not linked closely to the value of effort. Also, we assume an increasing-marginal-cost function that does not link investment directly to the value of the rent. These properties give us more freedom to "assume away" the potential participation problem by manipulating the profit function and make the contest worth of participation. For example, we can assume that the losers also gain some value (since studying effort may contribute to future productivity) and increase the expected value of participation. We can also add a negative constant to the cost function (studying may be fun at the beginning) and lower the cost of participation.

In short, we do not feel too uneasy to ignore the participation problem in this paper. However, this does not mean that we think the participation problem is not relevant to this problem. In this paper we find that the size and the composition of the competition pool are very importance in the effort-inducing function of the exam system. And the participation condition is definitely related to the size and composition of the competition pool.

B2. The second order condition

The second order condition may cause problem only when the equilibrium level of winning probability for student in question is less than 50%. For the profit function we consider, $V \times P(\mu_i) - \mu_i^2 / 2A_i$, the second order condition for student i's $V \times P''(\mu_i) - 1/A_i \leq 0$. If in the equilibrium $P(\mu_i) \geq 1/2$, then, under both the AP and the RP system, $P''(\mu_i) \leq 0$ and thus the condition must be satisfied. (Under the AP system, $P(\mu_i) \geq 1/2$ means $\mu_i \geq \underline{y}$. If μ_i increases, the gap between μ_i and $\underline{y}$ would increase and the MPWP would decrease. Similarly, under the RP system, $P(\mu_i) \geq 1/2$ means student i's score distribution in on the right hand side of the opponents' score distribution. If μ_i increases, the gap between the score distributions would increase and the MPWP would decrease.)

The potential violation of the second order condition can be avoided by adding a "handicap" term to the cost function. To put it clearly, the potential violation can be avoided by rewriting the cost function as $V \times P(\mu_i) - \mu_i^2 / 2(A_i - c_0)$, where c_0 is a constant term that represents a handicap that lowers the effectiveness of the ability to save educational costs.

To briefly summarize, the second order condition is automatically satisfied in part of our analyses. And for the other part, we can make reasonable adjustments to avoid the problem without changing the structure of our model.

C. The derivation of the MPWP in the RP system

Follow footnote (13) in the text, we first let $\theta(\cdot)$ and $\Theta(\cdot)$ be the pdf and cdf of ε. And again we have the property that $f_i(\hat{x}) = \theta(\hat{x} - \mu_i)$, $F_i(\hat{x}) = \Theta(\hat{x} - \mu_i)$, $f_j(\hat{x}) = \theta(\hat{x} - \mu_j)$, and $F_j(\hat{x}) = \Theta(\hat{x} - \mu_j)$.

The MPWP of student i's effort is

$$\tfrac{\partial}{\partial \mu_i}\left[prob(y_i > y_j) \right] = \tfrac{\partial}{\partial \mu_i}\left[\int_{\hat{y}} prob(y_i > \hat{y}) \times f_j(\hat{y}) d\hat{y} \right] = \tfrac{\partial}{\partial \mu_i}\left\{ \int_{\hat{y}} \left[1 - F_i(\hat{y})\right] \times f_j(\hat{y}) d\hat{y} \right\}$$

$$= \int_{\hat{y}} \left\{ \tfrac{\partial}{\partial \mu_i}\left[1 - F_i(\hat{y})\right] \right\} \times f_j(\hat{y}) d\hat{y} = \int_{\hat{y}} \left\{ \tfrac{\partial}{\partial \mu_i}\left[1 - \Theta(\hat{y} - \mu_i)\right] \right\} \times f_j(\hat{y}) d\hat{y} = \int_{\hat{y}} \theta(\hat{y} - \mu_i) \times f_j(\hat{y}) d\hat{y}$$

$$= \int_{\hat{y}} f_i(\hat{y}) \times f_j(\hat{y}) d\hat{y}$$

➔ the average value of the pdf of student i's score distribution using the pdf of student j's score distribution as the weights.

D. The derivation of the homogeneous 2-for-4 RP system

In the 2-for-4 homogeneous case, student i has to beat the second best of the other three students to get a position. The score distribution of his opponent can be found using the following order-statistic method.

Let's start with the notations. In the equilibrium, we expect every one has the same effort level and thus the same score distribution. Follow the text, we use $f(\cdot)$ to denote the pdf of their score distribution and $F(\cdot)$ the cdf. Student i's imaginary opponent's score distribution consists of the middle value of the 3 draws from the other 3 students' score distribution. And, once again, like in the text, we use $g(\cdot)$ and $G(\cdot)$ to denote the pdf and cdf of the opponent's score distribution.

Actually, in the homogeneous case, we can simply apply the well-known order-statistics formula to find the function of $g(\cdot)$. (For example, see David (1981).) The k^{th} highest value of the n draws from the same distribution characterized by $f(\cdot)$ and $F(\cdot)$ is $\frac{n!}{(n-k)!(k-1)!} \times F^{(n-k)} \times (1-F)^{(k-1)} \times f$. In this case, we simply plug in $n=3$ and $k=2$ into the formula, and get $g = \frac{3!}{(3-2)!(2-1)!} \times F^{(3-2)} \times (1-F)^{(2-1)} \times f = 6 \times F \times (1-F) \times f$.

In follows, we will derive this result from sketch to illustrate how the formula comes to live. We need to first find $G(\cdot)$ first and then derive $g(\cdot)$.

$G(x) =$ prob(the middle value of 3 draws is smaller or equal to x)
= prob(all of 3 draws are smaller or equal to x) + prob(2 of 3 draws are smaller or equal to x)
$= C_3^3 \times F(x)^3 + C_2^3 \times F(x)^2 \times [1-F(x)] = -2 \times F(x)^3 + 3 \times F(x)^2$
➔ $g(x) = \frac{\partial}{\partial x} G(x) = -6 \times F(x)^2 \times f(x) + 6 \times F(x) \times f(x) = 6 \times F(x) \times \left[1-F(x)\right] \times f(x)$
The result is identical to the one we got by applying formula.

To summarize, we derive the MPWP for student i 's effort is the 2-for-4 RP system as $\int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) g(\hat{y}) d\hat{y} = \int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) \times 6 \times F(\hat{y}) \times \left[1-F(\hat{y})\right] \times f(\hat{y}) d\hat{y}$.

E. The homogeneous case in general

E1. The 75%-admission-rate case (homogeneous)

- For the 3-for-4 case, the imaginary opponent of student i is the third best student of the other 3, and the opponent's score distribution is $g_4 = \frac{3!}{2!} \times F(x)^0 \times \left[1-F(x)\right]^2 \times f$
- For the 6-for-8 case, the imaginary opponent of student i is the 6th best student of the other 7, and the opponent's score distribution is $g_8 = \frac{7!}{1!5!} \times F(x)^1 \times \left[1-F(x)\right]^5 \times f$

…
- For the 30-for-40 case, the imaginary opponent of student i is the 30th best student of the other 39, and the opponent's score distribution is $g_{40} = \frac{39!}{9!29!} \times F(x)^9 \times \left[1-F(x)\right]^{29} \times f$

…
- For the 300-for-400 case, the imaginary opponent of student i is the 300th best student of the other 399, and the opponent's score distribution is $g_{400} = \frac{399!}{99!299!} \times F(x)^{99} \times \left[1-F(x)\right]^{299} \times f$

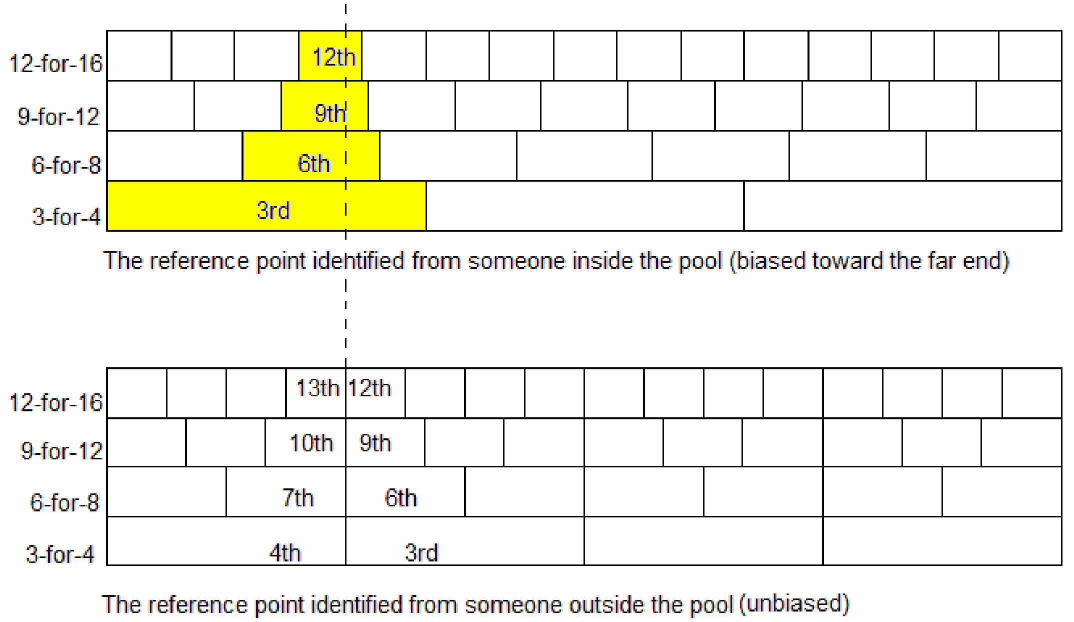
…
- For the 3000-for-4000 case, the imaginary opponent of student i is the 3000th best student of the other 3999, and the opponent's score distribution is $g_{4000} = \frac{3999!}{999!2999!} \times F(x)^{999} \times \left[1-F(x)\right]^{2999} \times f$
- The simulation result (deriving from a logistic distribution with variance σ^2)

	RP ($N=4$)	RP ($N=8$)	…	RP ($N=40$)	…	RP ($N=400$)	…	RP ($N=1000$)	…	AP
MPWP ($1/\sigma$)	0.27207	0.3023	…	0.331973	…	0.339239	…	0.339748	…	0.340087

(The MPWP under the AP system is the pdf measured at the 25% percentile, under the 3-for-4 RP system is $\int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) g_4(\hat{y}) d\hat{y}$, under the 6-for-8 RP system is $\int_{\hat{y}=-\infty}^{\infty} f(\hat{y}) g_8(\hat{y}) d\hat{y}$, and so on.)

- From the simulated result one can easily see that the MPWP under the 3-for-4 RP system is smaller than that under the AP system, that the MPWPs under the 75% RP system increases as N increases, and that the MPWPs under the 75% RP system approaches that under the AP system from below as $N \to \infty$.

E2. The graphic demonstration of the identification bias (for the 75% admission rate)



The bottom panel shows the "unbiased" identification of the imaginary opponent that represents the 25% of the competition pool. This opponent is the one between the 3rd and the 4th in the 3-for-4 case, the one between the 6th and the 7th in the 6-for-8 case, and so on. The identification of the imaginary opponent used in our RP model is shown in the top panel. The imaginary opponent is the 3rd of the other 3 students in the 3-for-4 case, the 6th of the 7 other students in the 6-for-8 case, and so one. We highlight this identification, and one can easily see that this identification is not unbiased around the 25% position: it biases toward the far end. As the competition pool expands, however, this bias becomes less significant.

We also find that this identification bias is negligible when the admission rate is close to 50% and is more significant when the admission rate is farther away from 50%. For an admission rate exactly equals 50%, this bias actually does not exist.

E3. The homogeneous case in general (the proof of the Result 1)

We take a random variable $x \sim \text{Logistic}\left(\mu, s\right)$, where μ and s are the mean and the scale parameter for the logistic distribution. (The variance of the distribution is actually $\frac{1}{3}\pi^2 s^2$.) . Let $f\left(\cdot\right)$ and $F\left(\cdot\right)$ be the pdf and cdf of x respectively. Their function forms are as the follows:

$$f\left(x\right) = \left[1 + \exp\left(-\frac{x-\mu}{s}\right)\right]^{-2} \times \frac{1}{s} \times \exp\left(-\frac{x-\mu}{s}\right), \text{ and}$$

$$F\left(x\right) = \left[1 + \exp\left(-\frac{x-\mu}{s}\right)\right]^{-1}$$

We consider an arbitrary rate of admission b/a, where both a and b are natural numbers, and $a > b$. The MPWP under the AP system is

$$f\left[F^{-1}\left(1 - b/a\right)\right] = f\left[\mu - s \times \log\left(-1 + \frac{1}{1 - b/a}\right)\right] = \frac{b \times \left(a-b\right)}{a^2 \times s} \qquad \text{(AD-1)}$$

And the MPWP under the RP system is $\int_{\hat{x}=-\infty}^{\infty} f(\hat{x}) \times g(\hat{x})\,d\hat{x}$, where

$g(x) = \frac{(a-1)!}{(a-b-1)!(b-1)!} \times F(x)^{(a-b-1)} \times \left[1-F(x)\right]^{(b-1)} \times f(x)$ is the pdf of the score distribution of the imaginary opponent: the b^{th} best student of the other $(a-1)$ students. We can solve for the MPWP under the RP system:

$$\int_{\hat{x}=-\infty}^{\infty} f(\hat{x}) \times g(\hat{x})\,d\hat{x} = \frac{b \times (a-b)}{(a+a^2) \times s} \tag{AD-2}$$

With (AD-1) and (AD-2), we can now prove the first point by comparing the MPWP under the AP and the MPWP under the RP system:

$$\frac{b \times (a-b)}{a^2 \times s} - \frac{b \times (a-b)}{(a+a^2) \times s} = \frac{b \times (a-b)}{s} \times \left[\frac{1}{a^2} - \frac{1}{(a+a^2)}\right] = \frac{b \times (a-b)}{s} \times \left[\frac{1}{a^2+a^3}\right] > 0 .\ ■$$

For the second point, we first set b/a equal a constant κ, $\kappa < 1$. We then rewrite the MPWP under the RP system as

$$\frac{b \times (a-b)}{(a+a^2) \times s} = \frac{\frac{b}{a} \times (1-\frac{b}{a})}{(\frac{1}{a}+1) \times s} = \frac{\kappa \times (1-\kappa)}{(\frac{1}{a}+1) \times s}$$

We can thus prove the second point by differentiating the MPWP under the RP system with the number of student a: $\frac{\partial}{\partial a}\left(\text{MPWP under the RP sustem}\right)\Big|_{b/a=\kappa} = \frac{\partial}{\partial a}\left[\frac{\kappa \times (1-\kappa)}{(\frac{1}{a}+1) \times s}\right] = \frac{\kappa \times (1-\kappa) \times \frac{1}{a^2}}{(\frac{1}{a}+1)^2 \times s} > 0$.■

For the third point, fixing an admission rate κ, the difference between the MPWP under the AP system and the MPWP under the RP system is $\frac{\kappa \times (1-\kappa)}{s \times (1+a)}$, and it is clear that $\frac{\kappa \times (1-\kappa)}{s \times (1+a)} \to 0$ from above as $a \to \infty$.■

E4. The case that the measurement error is assumed to be normal

In this part, we attempt to demonstrate that the three results shown in **Result 1** also hold for the case that the measurement error belongs to the normal distribution. We are not able to prove the results formally, but we would like to demonstrate that the results hold for all possible admission rates: we consider some representative rates and show that the results hold for all of them. Since the MPWP is symmetric around the 50% rate, we only consider the admission rate lower than 50%.

First of all, we show that for all those admission rates, the MPWP under the RP systems with the smallest scale (with one student admitted) is smaller than that under the corresponding AP system. One can easily see this result by comparing the second row and the fifth row of the following table.

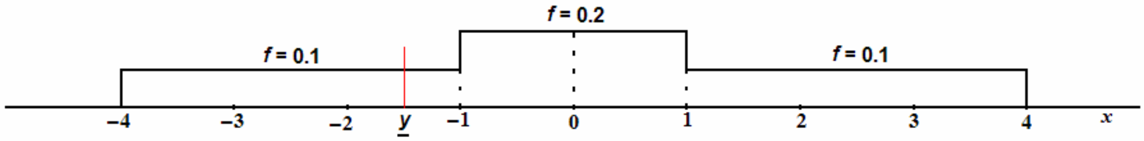
Admission Rate	1/2	1/3	1/4	1/5	1/6	1/10	1/50	1/100
MPWP (RP, $n=1$)	0.282095	0.282095	0.257344	0.232593	0.211201	0.153875	0.044982	0.025076
MPWP (RP, $n=2$)	0.331597	0.31816	0.284478	0.254011	0.228747	0.163754	0.046557	0.025799
MPWP (RP, $n=100$)	0.39738	0.362583	0.317041	0.279392	0.249389	0.175242	0.048378	0.026634
MPWP (AP)	0.398942	0.3636	0.317777	0.279962	0.249851	0.175498	0.048418	0.026652

(n is the number of positions, and the unit of the MPWPs is $1/\sigma$.)

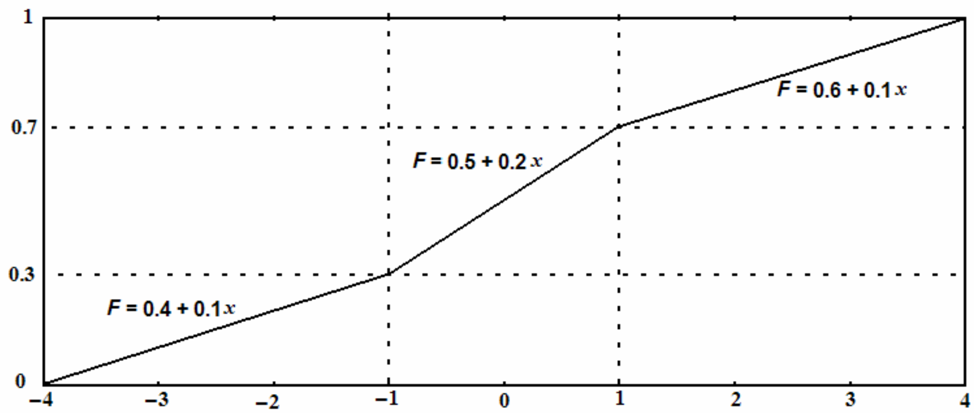
Secondly, we show that the MPWP under these RP systems increases as N increases: we increase the scale of each of the RP systems by adding one additional position. This result can be seen from the second row and the third row. And, thirdly, we show that these MPWPs approach their corresponding AP MPWPs when $N \to \infty$ by increasing the scale of the RP systems dramatically. By increasing the number of the positions to 100, one can see from the last two rows that the MPWPs under the RP systems are almost identical with those under the AP system.

E5. A counterexample (an example that the MPWP under the RP system is greater than the MPWP under the RP system for a 75% admission rate)

Consider a score distribution represented by the following figure:



The following figure shows the cdf of the distribution.



For the AP system, the MPWP is $f\left(\underline{y}\right)$ where $\underline{y}$ is set so that $F\left(\underline{y}\right)=25\%$. One can see that $f\left(\underline{y}\right)=0.1$.

For one of the students in the RP system, student i, his imaginary opponent is the worst student of the other 3 students. The pdf of his opponent's score distribution can be found by $g(\cdot)=\frac{3!}{2!}(1-F(\cdot))^2 f(\cdot)$.

And the MPWP of student i's effort is $\int_{\hat{y}=-4}^{4} f(\hat{y})g(\hat{y})d\hat{y}$. It can be shown that

$$\int_{\hat{y}=-4}^{4} f(\hat{y})g(\hat{y})d\hat{y}=$$

$$\int_{\hat{y}=-4}^{-1} 0.1\cdot\left[3\cdot\left(1-\left(0.4+0.1\hat{y}\right)\right)^2\cdot 0.1\right]d\hat{y}+\int_{\hat{y}=-1}^{1} 0.2\cdot\left[3\cdot\left(1-\left(0.5+0.2\hat{y}\right)\right)^2\cdot 0.2\right]d\hat{y}+\int_{\hat{y}=1}^{4} 0.1\cdot\left[3\cdot\left(1-\left(0.6+0.1\hat{y}\right)\right)^2\cdot 0.1\right]d\hat{y}=0.1316$$

In this case, the MPWP under the RP system is greater than the MPWP under the AP system. The intuition behind this is that the pdf of the score distribution drops dramatically around the criterion and gives the AP system a small MPWP. The MPWP under the RP system is an average value and is not affected by the dramatic decrease as much. The identification bias that we use to explain why the MPWP under the AP system is greater than that under the RP system still exists, but the effect of the dramatic fall dominates in this example.

F. The heterogeneous case with complete information

F1. The existence and the uniqueness of the solutions

A. the AP system

The solution under the AP system in this case consists of three equations: the first order condition for A_H, $V \cdot f_H(\underline{y}) = \mu_H / A_H$, the first order condition for A_L, $V \cdot f_L(\underline{y}) = \mu_L / A_L$, and the overall 50%-admission-rate condition $f_H(\underline{y}) = f_L(\underline{y})$.

Using the third condition, we can put the first two equations together and get

$$V \cdot f_H(\underline{y}) = V \cdot f_L(\underline{y}) = \mu_H / A_H = \mu_L / A_L$$

We start with an arbitrary μ_L. For this μ_L, we can find a unique μ_H that satisfies $\mu_H / A_H = \mu_L / A_L$. Namely, $\mu_H = \mu_L \times A_H / A_L$. This pair of μ_L and μ_H gives us a unique value of $f_H(\underline{y}) = f_L(\underline{y})$ ($\underline{y}$ must be in the middle of the two to bring an overall admission rate of 50%).

If we have the result that $V \cdot f_L(\underline{y}) > \mu_L / A_L$ (also $V \cdot f_H(\underline{y}) > \mu_H / A_H$), then we can raise the value of μ_L. Raising μ_L causes an even larger increase in μ_H and thus a larger gap between μ_L and μ_H. A larger gap means smaller value of $f_L(\underline{y})$. In short, we start with a situation that the LHS of A_L 's F.O.C. is greater than the RHS. And we can always balance the equation by increasing the RHS since doing so also decreasing the LHS. Sooner or later, the two sides will meet and we have a solution. Similarly, if we start with a larger RHS, then we could lower the value of μ_L to find a solution.

Once a solution is found, it must be unique. Since if we change the value of μ_L from the solution, the RHS and the LHS would always move toward opposite directions and the two sides would never meet again.

B. the 1-for-2 RP system

The solution under the 1-for-2 RP system in this case consists of two equations: the first order condition for A_H, $V \cdot \int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_L(\hat{y}) d\hat{y} = \mu_H / A_H$, and the first order condition for A_L, $V \cdot \int_{\hat{y}=-\infty}^{\infty} f_L(\hat{y}) f_H(\hat{y}) d\hat{y} = \mu_L / A_L$.

We can put the first two equations together and get

$$V \cdot \int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_L(\hat{y}) d\hat{y} = V \cdot \int_{\hat{y}=-\infty}^{\infty} f_L(\hat{y}) f_H(\hat{y}) d\hat{y} = \mu_H / A_H = \mu_L / A_L .$$

We start with an arbitrary μ_L. For this μ_L, we can find a unique μ_H that satisfies $\mu_H / A_H = \mu_L / A_L$. Namely, $\mu_H = \mu_L \times A_H / A_L$. This pair of μ_L and μ_H gives us a unique value of

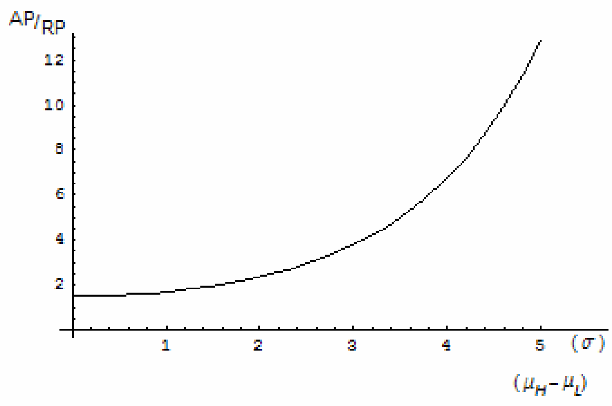
$$\int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_L(\hat{y}) d\hat{y} = \int_{\hat{y}=-\infty}^{\infty} f_L(\hat{y}) f_H(\hat{y}) d\hat{y} .$$

If we have the result that $V \cdot \int_{\hat{y}=-\infty}^{\infty} f_L(\hat{y}) f_H(\hat{y}) d\hat{y} > \mu_L / A_L$ (also $V \cdot \int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_L(\hat{y}) d\hat{y} > \mu_H / A_H$), then we can raise the value of μ_L. Raising μ_L causes an even larger increase in μ_H and thus a larger gap between μ_L and μ_H. A larger gap means smaller value of $\int_{\hat{y}=-\infty}^{\infty} f_L(\hat{y}) f_H(\hat{y}) d\hat{y}$. In short, we start with a situation that the LHS of A_L 's F.O.C. is greater than the RHS. And we can always balance the equation by increasing the RHS since doing so also decreasing the LHS. Sooner or later, the two sides will meet and we have a solution. Similarly, if we start with a larger RHS, then we could lower the value of μ_L to find a solution.

Once a solution is found, it must be unique. Since if we change the value of μ_L from the solution, the RHS and the LHS would always move toward opposite directions and the two sides would never meet again.

F2. The 1-for-2 case

Take two logistic distributions, *H* and *L*, to be the score distributions of A_H and A_L in the equilibrium. *H* is characterized by f_H, F_H, μ_H, and σ^2; *L* is characterized by f_L, F_L, μ_L, and σ^2. (They have the same variance, just different means.) Given these two score distributions, the MPWP generated by the AP system is $f_H(\frac{1}{2}\mu_H+\frac{1}{2}\mu_L)=f_L(\frac{1}{2}\mu_H+\frac{1}{2}\mu_L)$. The MPWP generated by the RP system is $\int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y})f_L(\hat{y})d\hat{y}$. We plot the ratio between these two MPWPs (AP over RP) in the following diagram. We can easily see that the ratio has a lower bound $3/2$ when $(\mu_H-\mu_L)\to 0$ and increases exponentially as $(\mu_H-\mu_L)$ increases. In short, the MPWP generated by the AP system is always greater than that by the RP system.



F3. The 2-for-4 case

Suppose that there are four students, 2 with A_H and the other 2 with A_L. For one of the students with A_H, his imaginary opponent is the second best student of the other 3, a group consisting of 1 A_H and 2 A_L. We name the pdf and cdf of this imaginary opponent's score distribution g_H and G_H. And we can derive g_H and G_H like the following.

$G_H(x)=$ prob(the middle value of 3 draws is smaller or equal to *x*)

= prob(all of 3 draws are smaller or equal to *x*) + prob(2 of 3 draws are smaller or equal to *x*)

= prob(all of 3 draws are smaller or equal to *x*)

+ prob(both draws from A_L are smaller or equal to *x* and the draw from A_H is greater than *x*)

+ prob(one draw from A_L is greater than *x*, the other draw from A_L and the draw from A_H are smaller or equal to *x*)

$$=C_1^1\times C_2^2\times F_H(x)^1\times F_L(x)^2+C_1^1\times C_2^2\times[1-F_H(x)]^1\times F_L(x)^2+C_1^1\times C_1^2\times F_H(x)^1\times F_L(x)^1\times[1-F_L(x)]^1$$

$$=-2\times F_H(x)^1\times F_L(x)^2+F_L(x)^2+2\times F_H(x)^1\times F_L(x)^1$$

➔

$$g_H(x)=\tfrac{\partial}{\partial x}G_H(x)$$

$$=-2\times F_L(x)^2\times f_H(x)-4\times F_H(x)\times F_L(x)\times f_L(x)+2\times F_L(x)\times f_L(x)+2\times F_L(x)\times f_H(x)+2\times F_H(x)\times f_L(x)$$

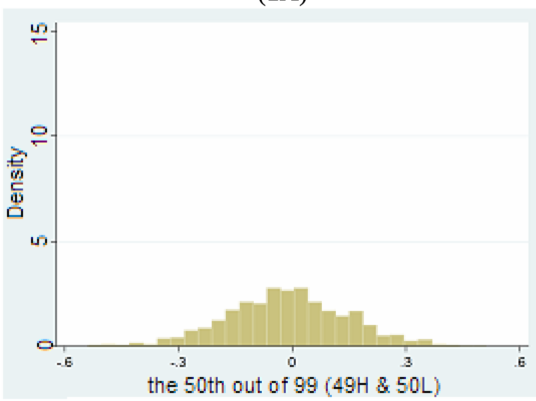
The MPWP for one of the student with A_H is $\int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y})g_H(\hat{y})d\hat{y}$. The MPWP for one of the student with A_L is $\int_{\hat{y}=-\infty}^{\infty} f_L(\hat{y})g_L(\hat{y})d\hat{y}$. And it is easy for one to see that these two MPWPs are identical.

F3. The case with a large N

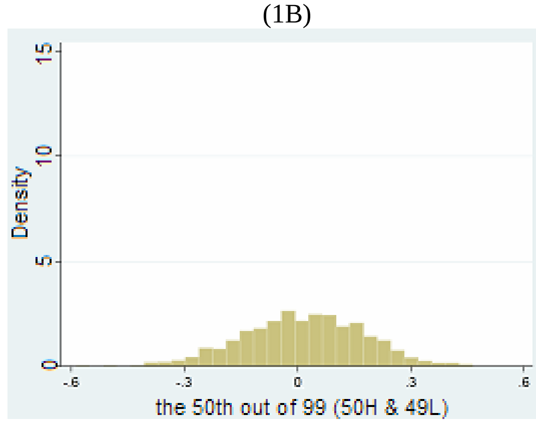
The analytical model gets out of hand (at least out of my hand) when N gets large. Thus we conduct some simulation to see what happens to the imaginary opponent's score distribution when N gets really large, as those shown in following figure. We take two logistic distributions with variance equal 1, one with mean 1 (the H distribution) and the other with mean -1 (the L distribution). To simulate the score distribution of the imaginary opponent faced by a student with high ability in a 50-for-100 test, we employ the following steps.

1. randomly draw 49 numbers from the H distribution and 50 numbers from the L distribution
2. mix these 99 numbers together and pick the 50^{th} large number (the middle number)
3. repeat step 1 and 2 for 1000 times
4. record these 1000 middle numbers and present the result using a histogram (1A)

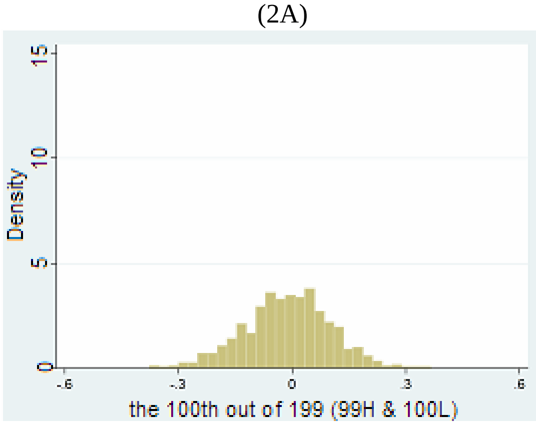
(1A)



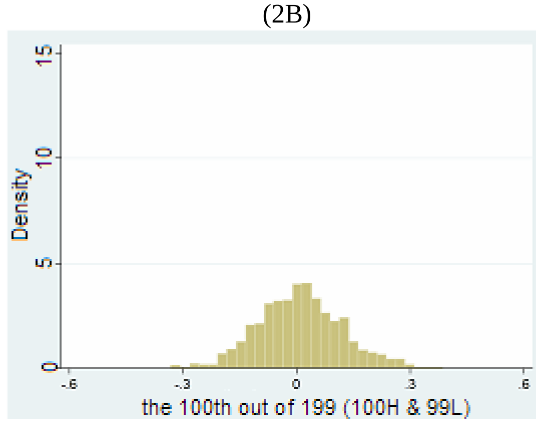
(1B)



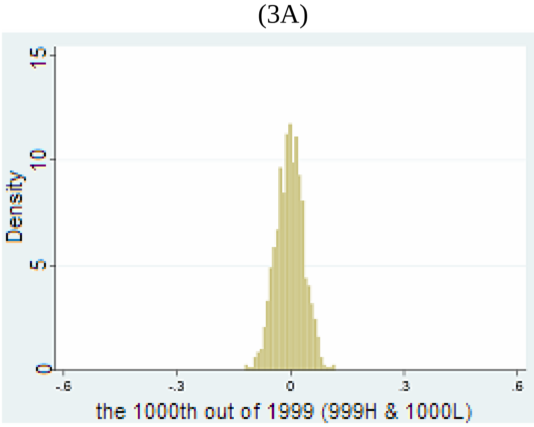
(2A)



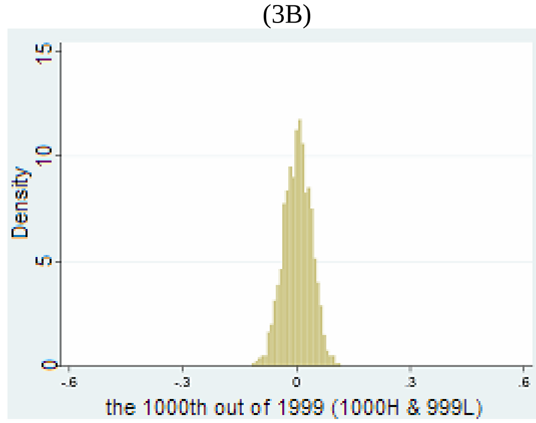
(2B)



(3A)



(3B)



Panel (1B) shows the score distribution of the imaginary opponent faced by a low-ability student. The only difference from the steps we use for (1A) is that, to get (1B), we draw 50 numbers from the H distribution and 49 from the L distribution.

The remaining panels are obtained by exactly the same method. (2A) draws 99 from the *H* distribution and 100 from the *L* distribution. (2B) draws 100 and 99, (3A) draws 999 and 1000, and (3B) draws 1000 and 999 from *H* and *L* respectively.

It is obvious that the score distributions collapse to a fixed number (0 in this case) as $N \to \infty$.

G. The heterogeneous case with incomplete information

G1. The 1-for-2 case

Consider a student with ability level A_H. He knows that there is a 50% chance that his opponent is of A_H and the other 50% chance to be of A_L. We assume students are risk neutral, thus the student's profit function is $\Pi = \frac{1}{2}\times V \times P_{HH} + \frac{1}{2}\times V \times P_{HL} - \mu_H^2 / 2A_H$, where P_{HH} is the student's chance of winning when he and his opponent both invest μ_H and P_{HL} is the student's chance of winning when he invests μ_H and his opponent invests μ_L. The first order condition is $\frac{1}{2}\times V \times P'_{HH} + \frac{1}{2}\times V \times P'_{HL} - \mu_H / A_H = 0$. P'_{HH} is the MPWP of the student's effort and it has the form $\int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_H(\hat{y}) d\hat{y}$; P'_{HL} is the MPWP of the student's effort and it has the form $\int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_L(\hat{y}) d\hat{y}$. The first order condition can be written as

$$\tfrac{1}{2}\times V \times \int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_H(\hat{y}) d\hat{y} + \tfrac{1}{2}\times V \times \int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) f_L(\hat{y}) d\hat{y} - \mu_H / A_H = 0,$$

which is identical to $V \times \int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) \left[\tfrac{1}{2} f_H(\hat{y}) + \tfrac{1}{2} f_L(\hat{y}) \right] d\hat{y} - \mu_H / A_H = 0$. In other words, the high-ability student in this case is like competing with a student who invests μ_H with a 50% chance and invests μ_L with the other 50%. And the MPWP of his effort is simply the average of the MPWP of his effort when he tries to beats these two different score distributions.

G2. The 2-for-4 case

Consider the case that there are 4 students, each of them has a 50% to have A_H and the other 50% to have A_L. For any student, there are 4 possible combinations for his opponents: he may face 3 A_H (with probability $1/8$), face 2 A_H and 1 A_L ($3/8$), face 1 A_H and 2 A_L ($3/8$), or face 3 A_L ($1/8$).

No matter which combination he faces, his imaginary opponent is the second best student of the other three. And the order statistic for these different combinations can be calculated using the method in Appendix F3:
If he faces 3 A_H, the pdf of his imaginary opponent's score distribution, g_{3H0L}, would be

$$6\times F_H(x)^2 \times f_H(x)$$

If he faces 2 A_H and 1 A_L, the pdf of his imaginary opponent's score distribution, g_{2H1L}, would be

$$-2\times F_H(x)^2 \times f_L(x) - 4\times F_L(x)\times F_H(x)\times f_H(x) + 2\times F_H(x)\times f_H(x) + 2\times F_H(x)\times f_L(x) + 2\times F_L(x)\times f_H(x).$$

If he faces 1 A_H and 2 A_L, the pdf of his imaginary opponent's score distribution, g_{1H2L}, would be

$$-2\times F_L(x)^2 \times f_H(x) - 4\times F_H(x)\times F_L(x)\times f_L(x) + 2\times F_L(x)\times f_L(x) + 2\times F_L(x)\times f_H(x) + 2\times F_H(x)\times f_L(x).$$

If he faces 3 A_L, , the pdf of his imaginary opponent's score distribution, g_{0H3L}, would be

$$6\times F_L(x)^2 \times f_L(x).$$

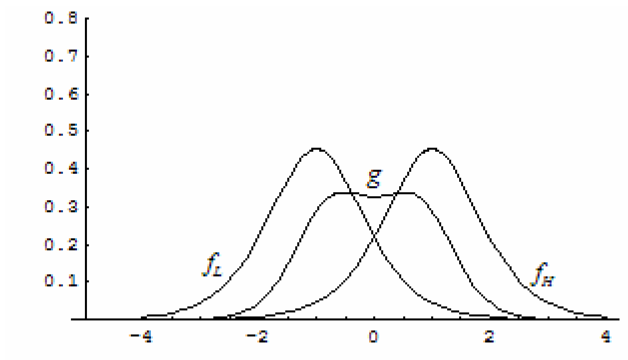
If the student we are considering has A_H, then the MPWP of his effort would be

$$\int_{\hat{y}=-\infty}^{\infty} f_H(\hat{y}) \left[\tfrac{1}{8} g_{3H0L}(\hat{y}) + \tfrac{3}{8} g_{2H1L}(\hat{y}) + \tfrac{3}{8} g_{1H2L}(\hat{y}) + \tfrac{1}{8} g_{0H3L}(\hat{y}) \right] d\hat{y}.$$

If he has A_L, then the MPWP of his effort would be

$$\int_{\hat{y}=-\infty}^{\infty} f_L(\hat{y})\left[\tfrac{1}{8}g_{3H0L}(\hat{y})+\tfrac{3}{8}g_{2H1L}(\hat{y})+\tfrac{3}{8}g_{1H2L}(\hat{y})+\tfrac{1}{8}g_{0H3L}(\hat{y})\right]d\hat{y}.$$

Because of symmetry, or one can easily see from the figure below, the average of the opponent's score distribution is symmetric around its mean $\frac{\mu_H+\mu_L}{2}$. And thus A_H and A_L have the same MPWP.



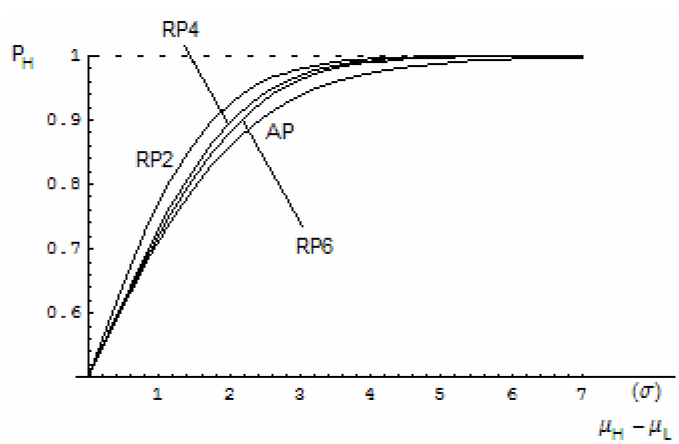
The figure above plots a logistic distribution with both mean and scale parameter equal to 1 (f_H), another logistic distribution with both mean equal to –1 and scale parameter equal to 1 (f_L), and the average of the opponent's score distribution calculated using the above analysis (g).

G3. When *N* is large

When the number of students in the competition pool increases, we expect the g distribution in the above figure to get more and more concentrated toward $\frac{\mu_H+\mu_L}{2}$. Unlike the complete information case, g distribution will not collapse into a number. This is because, in the structure we setup, the increase in N will not eliminate the possibilities that the compositions are different from the 50% A_H and 50% A_L composition.

H. The probability gap between the AP and the RP system with $N \geq 2$

The following graph is an expansion of < Figure 9 > in the text. The "RP4" curve represents the values of P_H in a 2-for-4 competition (2 A_H and 2 A_L compete for 2 positions) . And the "RP6" curve represents the values of P_H in a 3-for-6 competition (3 A_H and 3 A_L compete for 3 positions). One can easily see that the probability gap shrinks as N increases.



I. The "other" cases for more-position-more-effort problem

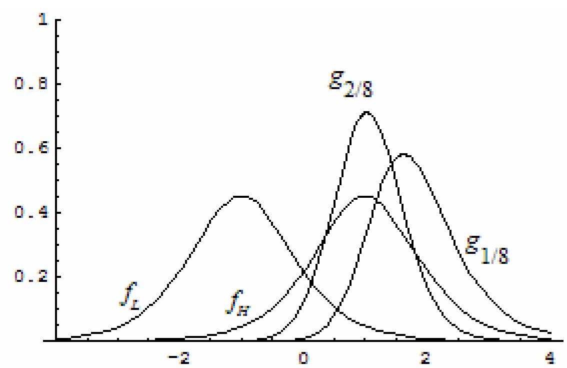
This part is to show that the phenomenon that an increase in admission rate causes students to study harder can also happen in the cases other than what are considered in the text.

I1. The complete information case

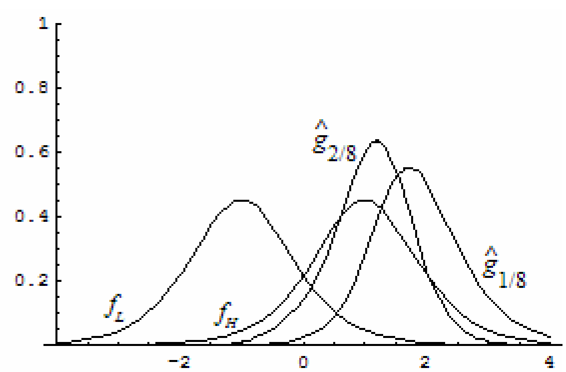
Consider the case that one more position is added to an initial competition that 4 A_H and 4 A_L compete for one position. For one of the students with A_H, he competes with 3 A_H and 4 A_L. When compete for just one position, he has to beat the best of the other 7 students. The score distribution of this imaginary opponent is $g_{1/8} = 3\times F_H^2 \times F_L^4 \times f_H + 4\times F_H^3 \times F_L^3 \times f_L$. When compete for two positions, he has to beat the second best of the other 7 students. The score distribution of this imaginary opponent is

$$g_{2/8} = -18\times F_H^2 \times F_L^4 \times f_H - 24\times F_H^3 \times F_L^3 \times f_L + 6\times F_H^1 \times F_L^4 \times f_H$$
$$+12\times F_H^2 \times F_L^3 \times f_L + 12\times F_H^2 \times F_L^3 \times f_H + 12\times F_H^3 \times F_L^2 \times f_L$$

We take 2 logistic distributions with scale parameter equal to 1, one with mean 1 (the *H* distribution) and the other -1 (the L distribution), to simulate the result. The relationship among the A_H student's score distribution, f_H, and the score distributions of his imaginary opponents in the 1-for-8 and 2-for-8 cases, $g_{1/8}$ and $g_{2/8}$, are plotted in the following figure (< Figure I1 >). It can be easily seen that $g_{2/8}$ is closer to f_H and thus generate a higher level of MPWP. In short, adding one more position make high-ability student work harder. The values of the MPWPs simulated by these two distributions are 0.268706 for the 1-for-8 case and 0.369068 for the 2-for-8 case. It can be shown easily that the low-ability students also work harder because of the additional position.



< Figure I1 >



<Figure I2>

I2. The incomplete information case

For the incomplete information case, we simply follow and modify the previous, the complete-information, case to illustrate the result.

For one of the students with A_H, he competes with different combinations of students with different probability: for instance, there is a $1/128$ chance that he would be competing with 7 A_H, a $35/128$ chance with 3 A_H and 4 A_L, and so on. The score distributions of for his imaginary opponent in different combinations and different admission rates can be calculated using similar method. We call the score distribution under the 1-for-8 case $\hat{g}_{1/8}$ and that under the 2-for-8 case $\hat{g}_{2/8}$. The similar relationship among f_H, f_L, $\hat{g}_{1/8}$ and $\hat{g}_{2/8}$ are plotted in the above figure (< Figure I2 >). Again, it can be easily seen that $g_{2/8}$ is closer to f_H and thus generate a higher level of MPWP. In short, adding one more position make high-ability student work harder. The values of the MPWPs simulated by these two distributions are 0.255183 for the 1-for-8 case and 0.347697 for the 2-for-8 case. One can easily see that the low-ability students also work harder because of the additional position.