American Kids, Why Don't They Study?

Sanghoon Lee*

University of Minnesota and Federal Reserve Bank of Minneapolis

February 13, 2004

Abstract

American students work less than Asian students in high school, but work more in college. We propose an explanation for this puzzle, using a two-stage-signaling model. Signaling can occur over time both in high school and college. We show that main signaling stage may be high school or college, and that students work harder in the main signaling stage. We also find that the main signaling is more likely to occur in high school if human networking is important for job productivity or if education environments among high school students are homogeneous.

Keywords: Education, Signaling, Many Signals, International Comparison.

JEL Classification: I21, J22, J24.

*Address: Department of Economics, University of Minnesota, 1035 Heller Hall, 271-19th Avenue South, Minneapolis, MN 55455; E-mail: sanghoon@econ.umn.edu. I am deeply indebted to Thomas Holmes and Andrea Moro for the encouragement and guidance. I also thank Hwikwon Ham, Oksana Leukhina, Sunghwan Kim, and Alice Schoonbroodt for valuable comments. Any remaining errors are mine.

1 Introduction

American high school students study substantially less than their East Asian counterparts and there is less pressure in high school. In East Asia it is not rare to see high school seniors staying in school from 7:30 am to 12:00 am every day. Some high school students even commit suicide because of their falling grades[1]. Obviously the pressure also exists in the US high school, but it is not of this magnitude. Walberg (2001) reports that Chinese students study, in and out of classroom, twice as much as American students, and Korean students study 83% more.

However, things are reversed in college. Asian students stop working when they get into college while American students start working diligently. A recent survey shows that average Japanese college and graduate school students study only 3 hours per day including class, which is shorter than the study time of their elementary school students (Japanese Statistics Bureau, 2001.) Average undergraduates at Seoul National University, the most exclusive college in Korea, study on average 12 hours per week outside classroom while Stanford University students study 26.1 hours and Ohio State University 20 hours. (Kim et al., 1999; Light and Wais, 2000; Lahmers and Zulauf, 2000.)

Why do American students work less than Asian students in high school, but work more in college? This paper proposes a signaling explanation of this phenomenon. It builds on the ideas of Spence (1973), and the main point of departure is that signaling can occur over time in both high school and college, and that societies can differ in when main signaling takes place. If the main signaling occurs in high school, students work hard in high school but do not work hard in college. If the main signaling occurs in college, students do not work hard in high school but work hard in college.

Our model generates two different kinds of explanations for why societies might have different signaling stages. The first explanation is a multiple equilibria argument. Our model shows that there exist multiple equilibria with different signaling stages under certain conditions. Thus, two societies with identical funda-

[1] The high pressure in East Asian high school has been described many times by the media. For example, refer to Elliot (1999), BBC (2000), and Gluck (2001).

mentals can have different signaling stages, and the main signaling stage is selected only by the society's self-fulfilling belief. The second explanation is based on the differences in fundamentals between societies. We show that main signaling is more likely to occur in high school if human networking is important for job productivity or if education environments are homogeneous among high school students. A case is made that these conditions are more true for East Asian countries than for the US.

This paper belongs to the theoretical literature on education performance and its determinants. Compared with the vast empirical literature, there is a relatively little theoretical literature in this field. Recent works include Lazear (2001) and Austen-Smith (2002), but each of them looks at aspects different from the one addressed here. There is also huge literature following the seminal work of Spence (1973), generalized in many ways including signaling with multiple signals[2]. What distinguishes this paper from the previous literature is that this paper focuses on the timing of signaling.

This theory has implications for two important issues. The first issue concerns the debate over the causes of the mediocre performance of American high school students. It is a well documented fact that the high school performance of American students is worse than that of their East Asian counterparts. For example, in a recent international student assessment American 15 year olds were ranked 14th in science while Koreans ranked 1st (OECD, 2000). While many factors may contribute to the poor performance of American students, undoubtedly one part of the explanation is simply that American high school students are not studying as hard as their East Asian counterparts.

This theory implies a trade-off between high school and college education performance, the levels of performance depending on when main signaling occurs. If the main signaling were to occur in high school as in East Asia, US students would work more in high school and their high school performance would improve, but they would work less in college and their college performance would decline. The mediocre performance of US high school education may then not be as bad as it looks, for it is one of the reasons that make US

[2]See Quinzii and Rochet (1985), Engers (1987), and Cho and Sobel (1990) for more information.

higher education performance so exceptional.

The second issue concerns the empirical literature estimating education productivity, that use international test data for high school students[3]. These studies compare the productivity of educational systems across countries. Any productivity study needs to control for all inputs, and in education one input is clearly how hard students are studying. However, any differences in main signaling stage leads to differences in study time and this would bias the estimates. For example, these studies conclude that education expenditure does not matter much for high school students performance. Part of what drives this result is that most East Asian countries belong to low spending group and their high school students do so well. If their excellent performance is at least partly due to main signaling occurring in high school, the importance of education expenditure will be underestimated.

2 The Model

In this section we present a model of students working in high school and college to signal their ability. There are three ability types of workers who are also heterogeneous in disutility from studying in high school. There are two colleges. One is considered the superior college, the other the inferior college. Three ability types and two colleges provide a minimal setting where signaling can take place both in high school and in college. Workers decide how much time to spend studying in high school and college, and which college to attend. There are two (or more) firms, and each of them maximizes its expected profit.

2.1 Workers

Workers differ in two characteristics – innate ability and disutility from studying in high school. There are three ability types represented by $\Theta \equiv \left\{\theta^H, \theta^M, \theta^L | \theta^H > \theta^M > \theta^L > 0\right\}$. Each ability type consists of a unit measure of workers heterogeneous in disutility coefficient γ of studying in high school. This disutility

[3] See Heyneman and Loxley (1983), Woessmann (2000), and Hanushek and Luke (2001) for more information.

coefficient γ is distributed, identically across all ability types, subject to a strictly increasing continuous cumulative distribution function $F : [\underline{\gamma}, \overline{\gamma}] \to [0,1]$. Every worker goes to high school and college, and the following utility function describes their preferences.

$$U_\gamma(n_h, n_c, w) \equiv \gamma v(n_h) + v(n_c) + w \text{ for all } \gamma \in [\underline{\gamma}, \overline{\gamma}] \text{ and } n_h, n_c, w \in \mathbb{R}_+$$

where n_h and n_c are time spent studying in high school and college respectively, and w is wage. The study utility function $v : \mathbb{R}_+ \to \mathbb{R}_-$ is twice differentiable, strictly *decreasing,* strictly concave, and satisfies $\lim_{n\to 1} v(n) = -\infty$ and $v(n) = -\infty$ for $n \geq 1$, which implies that no workers study more than a unit measure of time in either high school or college.

Note that a worker's disutility from studying in high school increases as the study disutility coefficient γ increases. The heterogeneity in γ captures other variances among students than innate ability. For example, γ would be high for those whose parents do not pay much attention to their children's education, or who just hate studying. In this model the heterogeneity in γ interferes with effective sorting in high school and allows an equilibrium where high ability workers with high γ end up in the inferior college and low ability workers with low γ in the superior college, making college name less informative of workers' innate ability. We assume for simplicity that there is no heterogeneity in disutility from studying in college.

2.2 Stage 1: High School

Each worker decides how much time to spend studying in high school. The high school performance p_h depends on the study time and the worker's ability. For simplicity we assume a linear performance function.

$$p_h(n_h, \theta) = \theta n_h \text{ for all } \theta \in \Theta \text{ and } n_h \in \mathbb{R}_+$$

2.3 Stage 2: Colleges

There are two colleges – A and B. Each college admits o*ne and half* unit measure of workers, using the cut-off rule based on workers' high school performances[4]. College A and B are *ex ante* identical, but *ex post* different in terms of the distribution of ability types, the study time of their students and their wage. Without loss of generality we assume that college A denotes the superior college in equilibrium, with the better average student ability.

Each worker applies for either college A or college B. Every worker ends up in one of the colleges because the total measure of admission from both colleges is equal to that of all workers. Once in college, workers decide how much time to spend studying. The college performance p_c is determined in the same manner as in high school.

$$p_c(n_c, \theta) = \theta n_c \text{ for all } \theta \in \Theta,\, n_h \in \mathbb{R}_+$$

The future job productivity of a worker grows by α fraction of the average students ability in the college he or she attends. In other words, the job productivity increases by $\alpha E(\theta|s)$ for a college s student $(s = A, B.)$ This productivity gain captures a "networking" effect on the job performance. College friends at work can help each other improving their job productivity. In addition, one's job productivity increases even more if his or her college friends have better abilities.

In East Asia, having a good network of people is crucial to one's success. Virtually every guide for doing business in China lists establishing the right "Guanxi" (connection) as one of the most important things to be successful[5]. In this model the networking effect coefficient α is a fundamental parameter deciding which type of equilibrium to exist. We show later that East Asian type equilibrium exists when α is big, and that US type equilibrium exists when α is small.

[4]In the equilibria we focus on this cut-off rule turns out to be the optimal strategy for each college maximizing its average students ability subject to the requirement that each college has to fill up their seats.

[5]For more information on "Guanxi" refer to Gold et al. (2002) or search in Google http://www.google.com.

2.4 Stage 3: Job Market

There are two firms (or more), indexed by 1 and 2, which maximize their expected profits. Each firm has the same CRS technology, where a job productivity of a worker is given by his or her innate ability plus the productivity gain in college, therefore a college s graduate with ability θ^i produces $\theta^i + \alpha E(\theta|s)$ ($s = A, B$ and $i = H, M, L$.) We assume that studying in either high school or college does not improve job productivity, and that there is no networking effect in high school. These unrealistic assumptions are made only for simplicity.

Firms compete for workers by simultaneously announcing their wage schedule. Firms can observe workers' high school performance, college name, and college performance. However, we focus on the equilibria where high school performance of a worker can be perfectly inferred from his or her college name and thus firms ignore high school performance. Therefore, firms' wage offers depend only on college name and college performance.

2.5 Equilibrium

A Bayesian Nash Equilibrium of this model consists of the list including each worker's study time and college choice, each college's cut-off level of high school performance for admission, and each firm's wage schedule, such that every player's strategy is the best response to the other players' strategies. We focus on the following two types of equilibria in order to show the differences in signaling stage and their effect on education performance.

Definition 1 *Asian equilibrium is a separating Bayesian Nash equilibrium, where college A has only high and medium ability workers and college B has only medium and low ability workers.*

Definition 2 *US equilibrium is a separating Bayesian Nash equilibrium, where each college has all three ability types of workers.*

It is clear from the equilibrium definitions that a college name is a better signal of a worker's ability in Asian equilibrium. In Asian equilibrium, the firms can safely infer that a college A graduate is at least of medium ability and a college B graduate is at most of medium ability. Thus, signaling in high school is stronger in Asian equilibrium because the college names are determined by high school performances.

We characterize equilibrium by going backward from the last stage. In the job market firms make wage offers based on workers' college name and college performance. The standard argument shows that in Bertrand competition for workers both firms will offer the same wage equal to workers' expected productivity.

Lemma 1 *Let $w_i(s, p_c) : \{A, B\} \times \mathbb{R}_+ \to \mathbb{R}_+$ be the equilibrium wage offer function of the firm i ($i = 1, 2$). Whenever there exists a positive measure of the workers who attend college s^* and perform p_c^* in the college, wage profile w_1 and w_2 satisfy*

$$w\left(s^*, p_c^*\right) \equiv w_1\left(s^*, p_c^*\right) = w_2\left(s^*, p_c^*\right) = E\left(\theta | s^*, p_c^*\right) + \alpha E\left(\theta | s^*\right).$$

Proof. All proofs are in the Appendix. ■

Lemma 1 implies that *every* worker, in both Asian and US equilibrium, will be paid his or her *true* productivity because both equilibria are *separating*.

3 Colleges

In this section we study how workers behave in each college. We find that college students work harder and perform better in US equilibrium and that the benefit of attending college A is greater in Asian equilibrium. We use the standard results from the Spence model to analyze signaling in colleges. The following lemma shows that the "single crossing" property holds between college performance and wage.

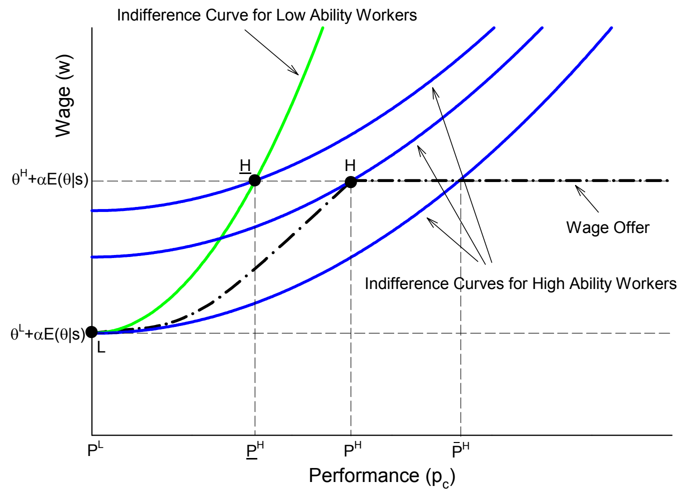

Figure 1: Spence Signaling Model

Lemma 2 *(Single Crossing Property)*

$$\frac{\partial}{\partial \theta}\left(-\frac{\partial \hat{U}/\partial p_c}{\partial \hat{U}/\partial w}\right) < 0$$

where $\hat{U}(\theta, p_c, w) \equiv v(p_c/\theta) + w$ *is the subutility function after high school.*

Figure 1 illustrates a simple application of the Spence signaling model to a hypothetical college s, where there exist only high and low ability workers. The single crossing property requires that the indifference curve of low ability workers is steeper than that of high ability workers wherever the two curves intersect. In any *separating* equilibrium the lowest ability workers do not study at all and get paid the wage equal to their true productivity. High ability workers can signal themselves away from low ability workers by attaining the level of performance P^H because it is not profitable for low ability workers to imitate.

However, the separating performance of high ability workers is not unique. In Figure 1 any performance between $\underline{P}^H$ and $\bar{P}^H$ can be supported as a separating equilibrium performance of high ability workers, with a suitable wage offer curve. Therefore, in order to obtain a unique prediction for college outcomes, we focus on the well known "Riley outcome." The Riley outcome is the unique Pareto-dominant separating equilibrium from Spence signaling model. It is also the only equilibrium that survives the refinement of D1 criterion[6]. In this paper we use an equivalent but more practical definition.

Definition 3 *Riley outcome for a college is the unique separating equilibrium within the college, where the lowest ability workers do not study at all, and the higher ability workers study just enough to achieve the minimum level of performance needed to weakly signal themselves away from the lower ability workers.*

For example, in Figure 1 P^L and $\underline{P}^H$ constitute the Riley outcomes for low and high ability workers respectively. Low ability workers do not work at all, therefore $P^L = 0$, and $\underline{P}^H$ is determined so that low ability workers are indifferent between their equilibrium outcome $L = \left(P^L, \theta^L + \alpha E\left(\theta|s\right)\right)$ and high ability workers' outcome $\underline{H} = \left(\underline{P}^H, \theta^H + \alpha E\left(\theta|s\right)\right)$ in Figure 1.

$$v(P^H/\theta^L) + \theta^H + \alpha E\left(\theta|s\right) = v(0) + \theta^L + \alpha E\left(\theta|s\right)$$

Canceling out $\alpha E\left(\theta|s\right)$ from both sides, we obtain

$$v(P^H/\theta^L) + \theta^H = v(0) + \theta^L.$$

Note that the college outcome P^L and $\underline{P}^H$ depend only on which ability types exist in the college. They are determined independently of high school disutility coefficient γ, networking coefficient α, or the share of each ability type within the college. This feature enables us to derive college outcomes only from the equilibrium definitions specifying which ability types exist in each college, and also turns out to be convenient when we prove the existence for each equilibrium.

[6] For more information on D1 criterion refinement, refer to Banks and Sobel(1987), Cho and Kreps(1987), and Cho and Sobel (1990).

3.1 Colleges in Asian Equilibrium

In Asian equilibrium college A has only high and medium ability types and college B has only medium and low ability types. The firms believe that college A graduates are at least of medium ability regardless of their college performance because no low ability workers exist in college A. In this separating equilibrium the medium ability workers in college A, the lowest ability type in this college, do not study at all $\left(P_A^M = 0\right)$ and get paid their productivity $\theta^M + \frac{\alpha}{3}\left(\theta^M + 2\theta^H\right)$, where $\frac{\alpha}{3}\left(\theta^M + 2\theta^H\right)$ is the productivity gain from the networking effect for college A graduates.

High ability workers study just enough to weakly separate themselves from the medium ability workers and get paid their true productivity $\theta^H + \frac{\alpha}{3}\left(\theta^M + 2\theta^H\right)$. Their equilibrium performance P_A^H is thus determined so that medium ability workers are indifferent between their equilibrium pay off $\left(0, \theta^M + \frac{\alpha}{3}\left(\theta^M + 2\theta^H\right)\right)$ and the high ability worker's pay off $\left(P_A^H, \theta^H + \frac{\alpha}{3}\left(\theta^M + 2\theta^H\right)\right)$.

$$v(P_A^H/\theta^M) + \theta^H = v(0) + \theta^M.$$

Even though college A does not have any low ability workers in equilibrium, we still need to know how the low ability workers would behave in college A because they compare college A and college B in high school when they decide which college to attend. A low ability worker who graduates from college A would be perceived at least of medium ability because the firms believe that any college A graduates are at least of medium ability. The low ability worker in college A would not choose to perform any better than medium ability workers, because it costs the low ability worker more effort to achieve the same level of performance. Since the medium ability workers in college A do not study at all, the low ability worker in college A would not study at all and get paid the wage of the medium ability workers who graduate from college A.

Figure 2 shows the indifference curves for each ability type in college A and an equilibrium belief of the firms represented by their wage offers. This figure illustrates how each ability type of workers and the firms

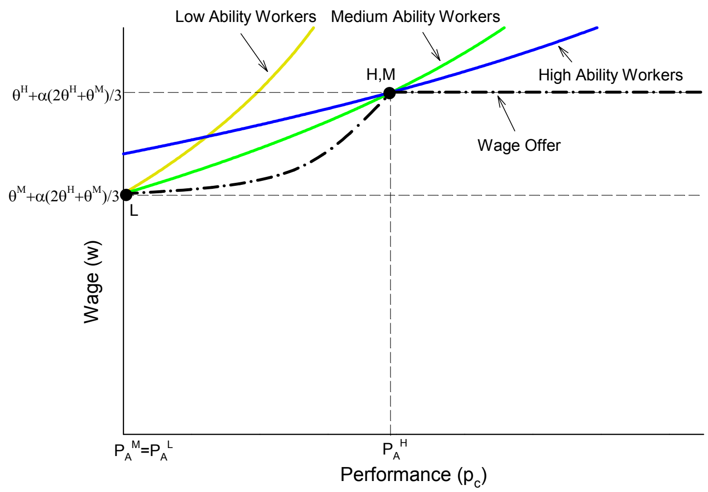

Figure 2: College A in Asian Equilibrium

optimize against the others' strategies in college A. Table 1 summarizes the equilibrium outcome.

Ability	Performance	Wage
Low	0	$\theta^M + \frac{\alpha}{3}\left(2\theta^H + \theta^M\right)$
Medium	0	$\theta^M + \frac{\alpha}{3}\left(2\theta^H + \theta^M\right)$
High	P_A^H such that $v(P_A^H/\theta^M) + \theta^H = v(0) + \theta^M$	$\theta^H + \frac{\alpha}{3}\left(2\theta^H + \theta^M\right)$

Table 1: College A in Asian Equilibrium

In college B, there are only low and medium ability workers. In this separating equilibrium, the low ability workers do not study at all and the medium ability workers study just enough to separate themselves from the low ability workers.

What would high ability workers do if they deviate to college B? The high ability workers would not earn more than medium ability workers in college B because firms believe that any college B graduate is at best of medium ability. Therefore, they do not perform any better than medium ability workers. The high ability

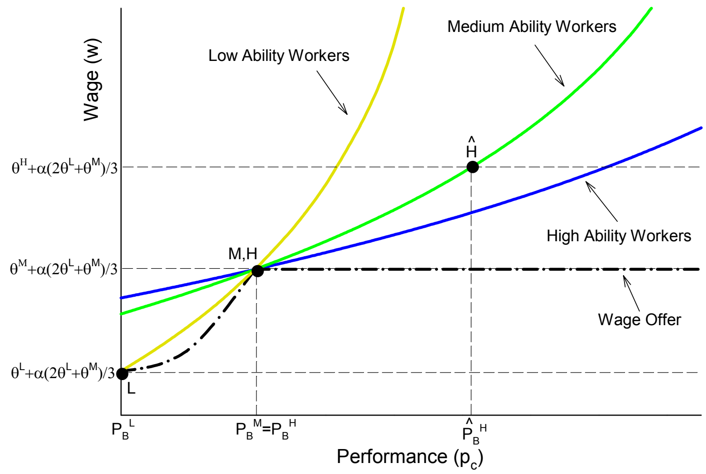

Figure 3: College B in Asian Equilibrium

worker would not perform any less either, because it is easier for the high ability worker to achieve the same level of performance. Figure 3 illustrates the equilibrium outcomes in college B, and Table 2 summarizes them.

Ability	Performance	Wage
Low	0	$\theta^L + \frac{\alpha}{3}\left(\theta^M + 2\theta^L\right)$
Medium	P_B^M such that $v(P_B^M/\theta^L) + \theta^M = v(0) + \theta^L$	$\theta^M + \frac{\alpha}{3}\left(\theta^M + 2\theta^L\right)$
High	$P_B^H = P_B^M$	$\theta^M + \frac{\alpha}{3}\left(\theta^M + 2\theta^L\right)$

Table 2: College B in Asian Equilibrium

3.2 Colleges in US Equilibrium

In US equilibrium each college has all three ability types. In each college low ability workers do not study at all, and the higher ability workers study just enough to separate themselves from the lower ability types. Figure 4 illustrates the college outcomes in US equilibrium and Table 3 summarizes them.

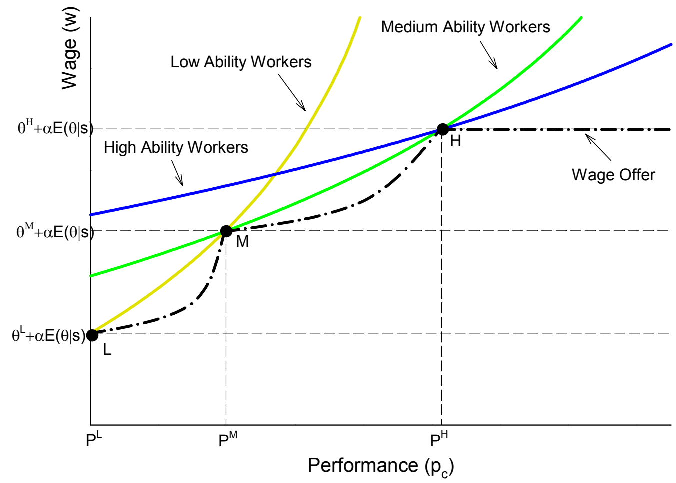

Figure 4: College s in US Equilibrium ($s = A, B$)

Ability	Performance	Wage
Low	0	$\theta^L + \alpha E\,(\theta\|s)$
Medium	P_s^M such that $v(P_s^M/\theta^L) + \theta^M = v(0) + \theta^L$	$\theta^M + \alpha E\,(\theta\|s)$
High	P_s^H such that $v(P_s^H/\theta^M) + \theta^H = v(P_s^M/\theta^M) + \theta^M$	$\theta^H + \alpha E\,(\theta\|s)$

Table 3: College s in US Equilibrium (s=A,B)

Note in Table 3 that college performances P_s^i are not affected by college name s ($i = H, M, L$ and $s = A, B$.) Therefore, the workers of same ability perform the same level in both colleges. However, wages for each ability type may be different across the colleges because $E\,(\theta|s)$ depends on the share of each ability types within the college as well as which ability types of workers exist.

3.3 Asian Colleges versus US Colleges - Performance and Study Time

In Asian equilibrium any college A graduate is considered at least of medium ability regardless of their college performance while in US equilibrium college A graduates can be of any ability type. This belief allows college A students in Asian equilibrium to study less in college to signal their ability, thus lowering their college performance. The following result can be easily obtained by comparing Table 1 and Table 2, with Table 3.

Proposition 1 *Every worker performs weakly better in college in US equilibrium than in Asian equilibrium. In particular, high and medium ability workers in college A in Asian equilibrium would perform strictly better in US equilibrium.*

It follows trivially that college students spend more time studying in US equilibrium because the performance is an increasing function of the study time.

Corollary 2 *Every worker studies weakly more in college in US equilibrium than in Asian equilibrium. In particular, high and medium ability workers in college A in Asian equilibrium would study strictly more in US equilibrium.*

3.4 The Benefit of Attending College A

In Asian equilibrium there are two endogenous effects that make college A preferred to college B. First, there is the "networking" effect. The productivity gain from college A is greater than that from college B because college A students have higher ability on average than college B students. Second, there is also a "sorting" effect which makes college A even more attractive. The sorting effect occurs because in Asian equilibrium the firms believe that college A graduates are at least of medium ability and college B graduates are at most of medium ability.

In order to better understand this sorting effect suppose that there is no networking effect ($\alpha = 0$.) Low ability workers prefer college A because they can make the medium ability workers' wage by attending college

A. Medium ability workers get the same wage whether they attend college A or college B (when $\alpha = 0$.) However, they prefer college A because they do not have to study at all in college A. High ability workers prefer college A because they get paid the wage of medium ability workers if they attend college B[7]. This sorting effect can be verified algebraically by comparing the outcomes in Table 1 and Table 2 with holding $\alpha = 0$.

In US equilibrium the sorting effect does not exist. The outcomes for both colleges are identical in Table 3 if there is no networking effect ($\alpha = 0$.) The networking effect still makes college A weakly preferred to college B because $E(\theta|A) \geq E(\theta|B)$ by assumption, but its size is smaller than the networking effect of Asian equilibrium because the average students ability difference between college A and B is maximized in Asian equilibrium. Thus the benefit of attending college A is greater in Asian equilibrium[8].

Proposition 3 *The benefit of attending college A is strictly greater in Asian equilibrium than in US equilibrium.*

4 High school

In this section we study the decision of workers in high school such as how much to study and which college to attend. We find that high school students perform better and tend to work harder in Asian equilibrium. We also find sufficient conditions for the existence of each equilibrium.

Two factors can affect which type of equilibrium exists. The first factor is the networking effect coefficient α. We find that an Asian equilibrium exists for α large enough, and that a US equilibrium exists for α small

[7] The sorting effect for the high ability type is a little more complicated than is explained in text because high ability workers' studying hours are different across colleges. It can be shown rigorously by the following argument. We show that high ability workers prefer college A pay off $\left(P_A^H, \theta^H\right)$ to college B pay off $\left(P_B^H, \theta^M\right)$ when $\alpha = 0$. *Medium* ability workers are indifferent between $\left(P_A^H, \theta^H\right)$ and $\left(0, \theta^M\right)$ in Figure 2, and indifferent between $\left(\hat{P}_B^H, \theta^H\right)$ and $\left(P_B^H, \theta^M\right)$ in Figure 3. Since $P_B^H > 0$, it follows that $\hat{P}_B^H > P_A^H$ and thus *high* ability workers prefer $\left(P_A^H, \theta^H\right)$ to $\left(\hat{P}_B^H, \theta^H\right)$. Since *high* ability workers prefer $\left(\hat{P}_B^H, \theta^H\right)$ to $\left(P_B^H, \theta^M\right)$ in Figure 3, they prefer $\left(P_A^H, \theta^H\right)$ to $\left(P_B^H, \theta^M\right)$.

[8] The benefit of attending college A is algebraically defined as $\left\{v\left(n_c^A\right) + w^A\right\} - \left\{v\left(n_c^B\right) + w^B\right\}$, where n_c^s is the amount of work in college s and w^s is the wage ($s = A, B$.)

enough. The second factor is the level of heterogeneity in high school study disutility γ. We show with a numerical simulation that Asian equilibrium is likely to exist if the heterogeneity in γ is small and that US equilibrium is likely to exist if the heterogeneity in γ is large. Interestingly, both Asian and US equilibrium can coexist with the same parameters. In this case of multiple equilibria, equilibrium is selected only by the society's self-fulfilling belief.

Since we focus on the equilibria where the firms ignore workers' high school performance and thus their wage offers do not depend on high school performance, the workers in high school perform just as much as the cut-off performance for the college they attend. The cut-off performance level C_B for college B admission is 0, because college B is less preferred and there are enough seats in colleges to accommodate all workers. Therefore, only the workers who attend college A would study in high school, working just enough to achieve the college A cut-off performance level C_A. Note that a worker's high school performance can now be perfectly inferred from his or her college name, which enables the firms to ignore the redundant signal of high school performance.

4.1 High School in Asian Equilibrium and the Existence of Asian Equilibrium

In Asian equilibrium college A has no low ability workers and college B has no high ability workers. In order to show the existence of Asian equilibrium, it suffices to show that no low ability workers in high school decide to attend college A and that no high ability workers in high school decide to attend college B, taking as given the Asian equilibrium college outcomes described in section 3.1.

In Asian equilibrium a half unit measure of medium ability workers attend college A and the other half unit measure of medium ability workers attend college B. More precisely, the medium ability workers with $\gamma < \gamma_m$, where $F(\gamma_m) = 0.5$, attend college A because they have a lower disutility of achieving the cut-off performance C_A for college A admission, and the medium ability worker with $\gamma > \gamma_m$ attend college B. The medium ability worker with γ_m is indifferent between the two colleges, which allows us to uniquely determine

C_A by solving

$$\gamma_m v\left(C_A/\theta^M\right) + v(0) + \theta^M + \frac{\alpha}{3}\left(2\theta^H + \theta^M\right) = \gamma_m v(0) + v\left(P_B^M/\theta^M\right) + \theta^M + \frac{\alpha}{3}\left(2\theta^L + \theta^M\right).$$

The above can be rewritten as

$$\gamma_m\left\{v(0) - v\left(C_A/\theta^M\right)\right\} = v(0) - v\left(P_B^M/\theta^M\right) + \frac{2\alpha}{3}\left(\theta^H - \theta^L\right). \tag{1}$$

The LHS of condition (1) is the net cost of attending college A to the medium ability worker with γ_m while the RHS is the net benefit of attending college A. In addition, the college A cut-off performance level C_A has to satisfy the incentive compatibility conditions for the other types. In order to show that no high ability worker deviates to college B, it suffices to show that the high ability worker with the highest disutility coefficient $\bar{\gamma}$ of studying in high school doesn't deviate to college B. Analogously, we need to show the low ability worker with $\underline{\gamma}$ doesn't deviate to college A.

Let $\underline{R}_A^L$ and $\bar{R}_A^H$ be the maximum high school performance levels which low ability workers with $\underline{\gamma}$ and high ability workers with $\bar{\gamma}$ are willing to achieve in order to attend college A. They are indifferent between attending college A with these reservation high school performances and attending college B with 0 high school performance. Therefore, $\underline{R}_A^L$ and $\bar{R}_A^H$ are determined by

$$\underline{\gamma}\left\{v(0) - v\left(\underline{R}_A^L/\theta^L\right)\right\} = \theta^M - \theta^L + \frac{2\alpha}{3}\left(\theta^H - \theta^L\right), \tag{2}$$

$$\bar{\gamma}\left\{v(0) - v\left(\bar{R}_A^H/\theta^H\right)\right\} = v\left(P_A^H/\theta^H\right) - v\left(P_B^M/\theta^H\right) + \theta^H - \theta^M + \frac{2\alpha}{3}\left(\theta^H - \theta^L\right). \tag{3}$$

It follows from conditions $(1), (2)$, and (3) that $\underline{R}_A^L$, C_A, and $\bar{R}_A^H$ converge to θ^L, θ^M, θ^H respectively as α increases to infinity (note that college outcomes P_A^H and P_B^M do not depend on α), which implies that $\underline{R}_A^L < C_A < \bar{R}_A^H$ for sufficiently large α. For these α no low ability workers attend college A and no high ability workers attend college B, and therefore an Asian equilibrium exists.

Proposition 4 *There exists an Asian equilibrium for sufficiently large α.*

The intuition behind Proposition 4 is that workers reveal their true ability when the benefit of attending college A is big enough. When the benefit is small, the lower ability workers with low γ may outperform the higher ability workers with high γ. However, as the benefit becomes bigger and bigger, all higher ability workers eventually outperform any lower ability workers because they have higher upper bound on their performance.

4.2 High School in US Equilibrium and the Existence of US Equilibrium

In US equilibrium both college A and college B have all three ability types. Heterogeneity in high school study disutility γ makes the existence of US equilibrium possible, where some high ability workers with high γ attend the inferior college B and some low ability workers with low γ attend the superior college A. Therefore, in order for a US equilibrium to exist, the heterogeneity in γ has to be sufficiently large relative to heterogeneity in ability.

Assumption 1 $\frac{\bar{\gamma}}{\gamma_m} > \frac{\theta^H}{\theta^M}$ *and* $\frac{\gamma_m}{\underline{\gamma}} > \frac{\theta^M}{\theta^L}$ *where* $F(\gamma_m) = 0.5$.

Unlike Asian equilibrium, the ability distribution of the workers across the colleges in US equilibrium is not directly pinned down by the equilibrium definition but has to be endogenously determined. Since the sorting effect is not present in US equilibrium, the networking effect constitutes the entire benefit of attending college A. The workers in high school observe the size of this networking effect and decide which college to attend, aggregately determining the ability distribution of the workers across colleges. This new ability distribution, in turn, determines the size of the new networking effect. In equilibrium, the initial networking effect has to coincide with the resulting networking effect.

When the networking effect coefficient $\alpha > 0$ is fixed, the networking effect is determined by the difference in average students ability between the colleges. This cross ability difference $x \equiv E(\theta|A) - E(\theta|B)$ can not be negative because we are assuming that college A students have better average ability. Further, x is smaller than $\frac{2}{3}\left(\theta^H - \theta^L\right)$ which can be achieved only in Asian equilibrium. Let $\Pi \equiv [0, \frac{2}{3}\left(\theta^H - \theta^L\right)]$ denote the set of possible ability difference and let $\psi(x) : \Pi \twoheadrightarrow \Pi$ be the new cross ability difference correspondence resulting from the workers' college choices, given the initial ability difference x[9].

Suppose $x = 0$. There is no networking effect and workers are indifferent between college A and college B. The cut-off performances for admission have to be the same across the colleges, and workers randomly choose their colleges. Thus, the new ability distribution of workers is not unique and the resulting ability difference can be any number in Π, which implies

$$\psi(0) = \Pi.$$

Suppose $x > 0$. The cut-off performance for college B admission C_B is 0 because college B is strictly less preferred and the total measure of college admission is equal to the total measure of workers. Given x and $C_A > 0$, there exists the unique critical disutility coefficient $\tilde{\gamma}^i \in \mathbb{R}_{++}$ for each ability type, such that the worker with $\tilde{\gamma}^i$ is indifferent between college A and college B ($i = H, M, L$).

$$\tilde{\gamma}^i \cdot v(C_A/\theta^i) + \alpha E(\theta|A) = \tilde{\gamma}^i \cdot v(0) + \alpha E(\theta|B) \quad \text{for } x > 0,\ i = H, M, L$$

Solving the above equation for $\tilde{\gamma}^i$ we obtain

$$\tilde{\gamma}^i(x, C_A) = \frac{\alpha\{E(\theta|A) - E(\theta|B)\}}{v(0) - v\left(C_A/\theta^i\right)} = \frac{\alpha x}{v(0) - v\left(C_A/\theta^i\right)} \quad \text{for } x, C_A > 0,\ i = H, M, L \tag{4}$$

For each ability type i, those workers with γ lower than $\tilde{\gamma}^i(x, C_A)$ attend college A and the others attend college B. Since college A admits one and a half unit measure of workers, the cut-off performance C_A for

[9] ψ is a correspondence because ψ has a set value at $x = 0$.

college A is uniquely determined by the following condition.

$$\sum_{i=H,M,L} F\left(\tilde{\gamma}^i\left(x, C_A\right)\right) = 1.5 \text{ for } x > 0 \tag{5}$$

The above condition implicitly defines C_A as a function of x. Therefore $\tilde{\gamma}^i$ also becomes a function of only x. Since $\tilde{\gamma}^i(x)$ $(i = H, M, L)$ pins down the unique ability distribution across the colleges, $\psi(x)$ is uniquely determined (thus single valued) for $x > 0$.

$$\begin{aligned} \psi(x) &\equiv E(\theta|A, x) - E(\theta|B, x) \\ &= \frac{1}{1.5}\sum_{i=H,M,L} \theta^i F\left(\tilde{\gamma}^i(x)\right) - \frac{1}{1.5}\sum_{i=H,M,L} \theta^i \left\{1 - F\left(\tilde{\gamma}^i(x)\right)\right\} \text{ for } x > 0. \end{aligned} \tag{6}$$

So far we have assumed that α is fixed when characterizing $\tilde{\gamma}^i(x)$ and $\psi(x)$. Since we want to relate α to the existence of US equilibrium, we slightly modify the notations in order to reflect the effect of a change in α on $\tilde{\gamma}^i(x)$ and $\psi(x)$. Let $\tilde{\gamma}^i(x; \alpha')$ and $\psi(x; \alpha')$ denote $\tilde{\gamma}^i(x)$ and $\psi(x)$ respectively with $\alpha = \alpha'$. The following lemma is crucial for analyzing the conditions determining the existence of US equilibrium.

Lemma 3 *Let $\underline{\gamma}$, $\bar{\gamma}$, θ^H, θ^M, θ^L and F satisfy Assumption 1.*

(i) There exists $\gamma_0^i \in \left(\underline{\gamma}, \bar{\gamma}\right)$ such that for all $x > 0$ $\lim_{\alpha \searrow 0} \tilde{\gamma}^i(x; \alpha) = \gamma_0^i$ $(i = H, M, L.)$

(ii) For all $\alpha \geq 0$, $\psi(x; \alpha)$ is continuous, weakly increasing in x.

(iii) There exists $y_0 \in \left(0, \frac{2}{3}\left(\theta^H - \theta^L\right)\right)$ such that (a) for all $\alpha \geq 0$ $\lim_{x \searrow 0} \psi(x; \alpha) = y_0$ (b) $\lim_{\alpha \searrow 0} \psi\left(\frac{2}{3}\left(\theta^H - \theta^L\right); \alpha\right) = y_0$.

The part (i) of Lemma 3 says that each college comes to have all three ability types (note that $\underline{\gamma} < \gamma_0^i < \bar{\gamma}$) as the networking coefficient α converges down to 0. The part (ii) and (iii) of Lemma 3, illustrated in Figure 5, characterize the graphical properties of $\psi(x; \alpha)$. The part (iii) of Lemma 3 says that (a) the y axis intercept y_0 of $\psi(x; \alpha)$ is the same regardless of α, that (b) $\psi\left(\frac{2}{3}\left(\theta^H - \theta^L\right); \alpha\right)$ converges to the same

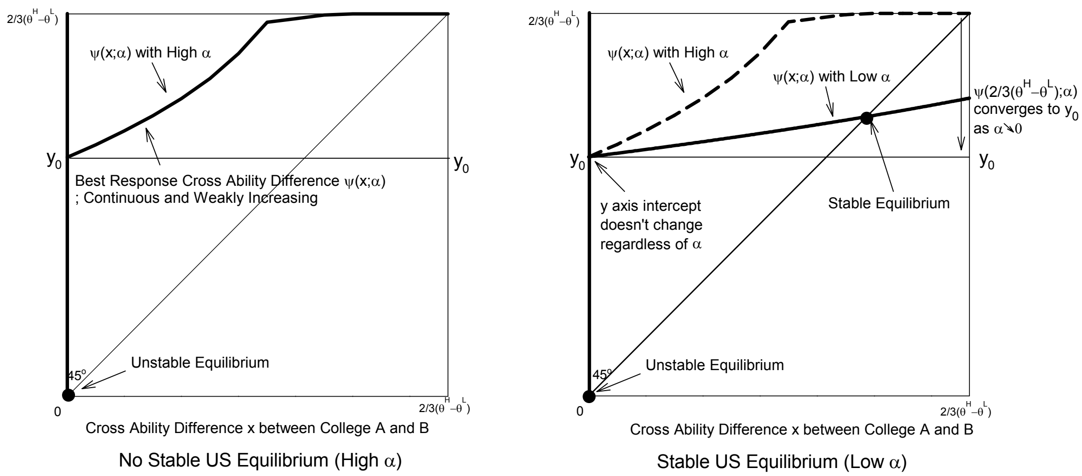

Figure 5: Existence of Stable US Equilibrium and the Networking Effect Coefficient α

y_0 as α converges down to 0, and that y_0 is strictly between 0 and $\frac{2}{3}\left(\theta^H - \theta^L\right)$.

In order for a US equilibrium to exist, the following two conditions have to be satisfied. First, each college has to have all three ability types. Second, the initial networking effect has to coincide with the new networking effect resulting from the workers' best response college decisions.

$$x^* \in \psi(x^*). \tag{7}$$

Suppose $x^* = 0$. Condition (7) is satisfied because $0 \in \Pi = \psi(0)$, and there exists a trivial US equilibrium where all workers are indifferent between both colleges and each worker flips a fair coin between the colleges. In this trivial US equilibrium, each college has the same share of all three ability types. However, this trivial US equilibrium is not stable defined in the following sense.

Definition 4 *Denote* $x_{n+1} \equiv \psi(x_n)$ *for all* $n \in \mathbb{N}$*. A US equilibrium with the cross ability difference* x^* *is (locally) stable if there exists* $\delta > 0$ *such that* $\lim_{n\to\infty} x_n = x^*$ *for all* $x_0 \in (x^* - \delta, x^* + \delta) \cap \Pi$*.*

Definition 4 is a usual definition of stability, that the system returns to the original equilibrium after small disturbances. The trivial US equilibrium at $x^* = 0$ is not stable, because y axis intercept y_0 of $\psi(x;\alpha)$ is strictly greater than 0 and $\psi(x;\alpha)$ is continuous and weakly increasing, due to the part (*ii*) and (*iii*) of Lemma 3.

Suppose $x^* > 0$. For $\alpha > 0$ sufficiently close to 0, there exists $x^* \in \left(0, \frac{2}{3}\left(\theta^H - \theta^L\right)\right)$ satisfying the condition (7) by the intermediate value theorem, because y axis intercept y_0 is strictly between 0 and $\frac{2}{3}\left(\theta^H - \theta^L\right)$, $\psi(x;\alpha)$ is continuous and weakly increasing in x, and $\psi\left(\frac{2}{3}\left(\theta^H - \theta^L\right);\alpha\right)$ converges down to y_0 as α converges down to 0. The following condition guarantees that each college has all three ability types.

$$\underline{\gamma} < \tilde{\gamma}^i(x^*;\alpha) < \bar{\gamma} \quad \text{for } i = H, M, L. \tag{8}$$

For $\alpha > 0$ sufficiently close to 0, condition (8) is also satisfied by the part (*i*) of Lemma 3. Therefore, for $\alpha > 0$ sufficiently close to 0, both condition (7) and (8) are satisfied and there exists a US equilibrium. US equilibrium may not be unique, but the stability condition is satisfied for at least one US equilibrium because there exists at least one x^* where ψ intersects the 45° line from above and ψ is weakly increasing.

Proposition 5 *Let $\underline{\gamma}$, $\bar{\gamma}$, θ^H, θ^M, θ^L and F satisfy Assumption 1. There exists a stable US equilibrium for $\alpha > 0$ sufficiently close to* 0.

The intuition behind Proposition 5 is the opposite to that of Proposition 4. Assumption 1 guarantees that low ability workers with $\underline{\gamma}$ outperform the high ability workers with $\bar{\gamma}$ when the benefit of attending college A becomes sufficiently small.

4.3 Why Asian Equilibrium in East Asia, and US Equilibrium in US?

The networking effect and the heterogeneity in γ plays an important role in deciding which type of equilibrium occurs. High networking effect implies high benefit of attending college A, which is likely to lead to Asian

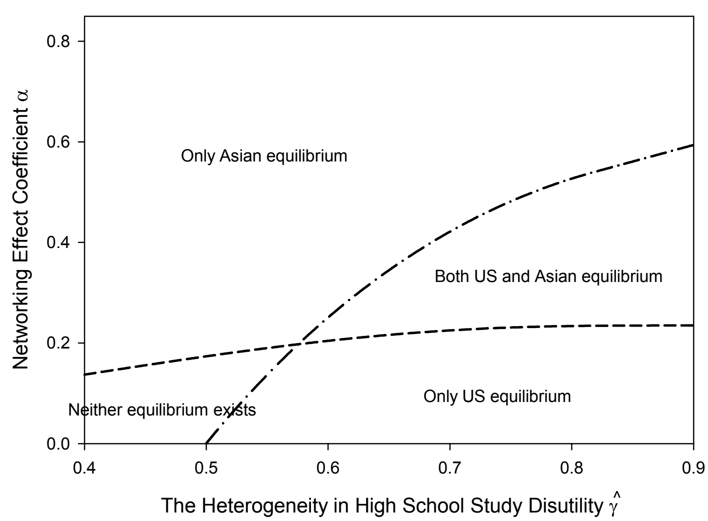

Figure 6: Networking Effect, Heterogeneity in γ, and the Existence of Equilibrium

equilibrium. High heterogeneity in high school study disutility γ interferes with ability sorting in high school, which is likely to lead to a US equilibrium.

Figure 6 summarizes the existence result for each type of equilibrium with different disutility coefficients α and levels of heterogeneity $\hat{\gamma}$, where $\theta^H = 3,\ \theta^M = 2,\ \theta^L = 1$, $U(\gamma, n_h, n_c, w) = \gamma \log(1-n_h) + \log(1-n_c) + w$ and $\gamma \sim$Uniform$[\underline{\gamma}, \bar{\gamma}]$ where $\underline{\gamma} = 1 - \hat{\gamma}$ and $\bar{\gamma} = 1 + \hat{\gamma}$.

Assumption 1 is satisfied for $\hat{\gamma} \geq 0.5$. Figure 6 shows that for this range of $\hat{\gamma}$ there exists a US equilibrium for sufficiently small α, as predicted by Proposition 5. It also shows that there exists an Asian equilibrium for sufficiently large α for each $\hat{\gamma}$, as predicted by Proposition 4. The range of α where an Asian equilibrium exists shrinks as the heterogeneity $\hat{\gamma}$ increases. In contrast, the range of α where a US equilibrium exists expands as the heterogeneity $\hat{\gamma}$ increases.

There are two different kinds of explanations, both consistent with this model, about why US equilibrium occurs in the US and why Asian equilibrium in East Asia. The first explanation applies to the region of the parameters in Figure 6, where both types of equilibria coexist. In this case both US and East Asia are interpreted as having the same parameters and the equilibrium is selected based only on the society's self-fulfilling belief. In East Asia, firms believe that a worker's college name is a relatively better indicator of ability than his or her college GPA. As a result, workers study hard in high school to get into a better college but do not study hard in college, which in turn makes the college name a better signal of ability than college GPA, fulfilling the firms' initial belief. In the US, firms believe that a worker's college GPA is a relatively better signal of ability than the college name. Consequently, workers do not study hard in high school but study hard in college, which in turn makes the college GPA a better signal of ability than a college name, thus fulfilling the firm's initial belief.

The second explanation is that East Asia and the US actually have different fundamental parameters, especially the networking effect coefficient α and the heterogeneity in high school study disutility γ. East Asia seems to have higher α. Human networking has been regarded as one of the most important things to be successful in East Asia. As mentioned, "Guanxi" (connection) is one of the most important things to be successful in China. In Japan and Korea, each college's graduates form their own clique and exchange favors among them. This exclusive favoritism is such a big social problem as to have its own name, "Gakubatsu" in Japan and "Hakbul" in Korea[10].

Heterogeneity in high school study disutility γ captures diversity in education environments among high school students. For example, the heterogeneity in γ is high if parents' interest over their children's education is diverse among population. It is not as obvious as the networking effect that the US has higher heterogeneity in γ, but it is possible that the tradition of Confucianism in East Asia, which emphasizes education, may have lead to more homogenous agreement on the value of education, resulting in the small variance in γ.

[10] For more information on academic clique in Japan and Korea, refer to pages 27-30 in Lafayette and Mente (1994) and Park (2003) respectively.

4.4 Performance and Study Time

The benefit of attending college A is greater in Asian equilibrium than in US equilibrium, when both equilibria coexist with the same parameters. Therefore, every worker in high school is willing to study more to attend college A in Asian equilibrium and their cut-off performance for college A admission is higher. Since high school students perform only as much as the cut-off performance of the colleges they attend, the average high school performance is better in Asian equilibrium.

Proposition 6 *Whenever both Asian and US equilibrium coexist with the same parameters, the average performance of high school students is strictly better in Asian equilibrium.*

Now we want to compare high school performance across Asian and US equilibrium with different networking effect coefficients α. Proposition 7 directly follows from Lemma 6 because the benefit of attending college A becomes even bigger as the networking effect coefficient α increases.

Corollary 7 *Suppose that an Asian equilibrium and a US equilibrium exist where α is weakly greater in the Asian equilibrium and the other parameters are the same. The average performance of high school students is strictly better in Asian equilibrium.*

These predictions are fairly weak because we expect that Asian equilibrium usually has a better high school performance even when Asian equilibrium has a lower α than US equilibrium. The extreme ability distribution across colleges in Asian equilibrium, which strengthen the networking effect, and the sorting effect that are found only in Asian equilibrium are often more than enough to compensate the loss in the benefit of attending college A resulting from low α. Figure 7 depicts high school performance and work amount in each equilibrium for different α's using the same parameters as used for Figure 6, except that $\hat{\gamma}$ is fixed at 0.8. For these parameters, all Asian equilibrium has higher high school performance, regardless of α, than any US equilibrium.

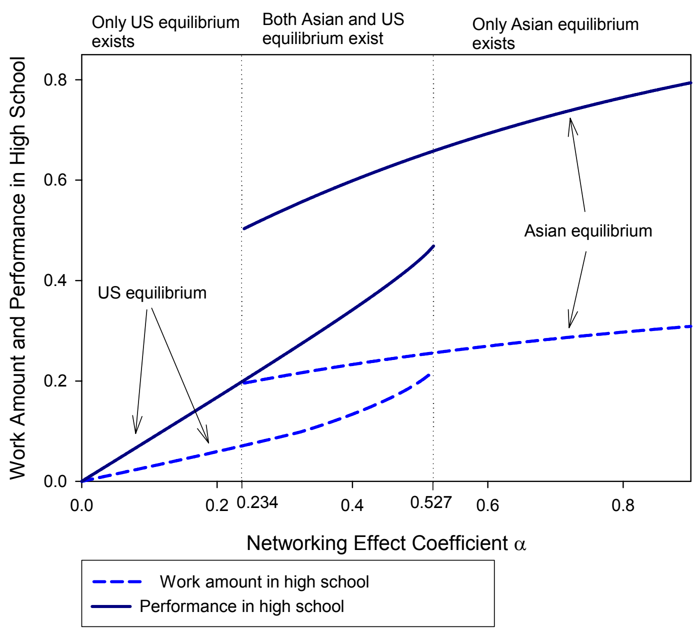

Figure 7: Study Time and Performance in High School

Rather surprisingly, it is not always true that workers in high school study more in Asian equilibrium even when they have better high school performance. In this model, only those attending college A study in high school in either equilibrium. Since the workers attending college A in Asian equilibrium has better ability on average than in US equilibrium, they need less study time to achieve the same level of performance. It is thus possible that high school students may work less in Asian equilibrium even when their performance is better. However, in most cases we expect high school students to work more in Asian equilibrium as in Figure 7.

5 Conclusion

Why do American students work less than Asian students in high school, but work more in college? In this paper we propose a signaling explanation for this puzzle. The main signaling stage occurs in college in US while it occurs in high school in East Asia. Therefore, when hiring workers, US firms weigh college GPA more seriously than college name determined by high school performance, while East Asian firms do the opposite. For this reason, US students work hard in college while East Asian students work hard in high school.

Our model generates two different kinds of explanations why societies might have different main signaling stages. The first explanation is that both US and East Asia may have the same fundamental parameters but each region is in a different equilibrium from the multiple equilibria. The second explanation is that US and East Asia may actually have different fundamentals. We show that main signaling stage is likely to be high school if human networking is important for job performance, or if the education environments, such as the parents' interest in their children's education, are homogeneous among high school students.

This theory has implication for the debate about the mediocre performance of US high school students. There is a trade off between high school and college education performances. If main signaling stage were to occur in high school as in East Asia, the US would have the better high school performance but its college performance would decline. The mediocre performance of US high school education may then not be as bad as it looks, for it is one of the reasons that make US higher education performance so exceptional.

This theory also has an important implication on recent empirical studies estimating the effect of different education systems on student performance, using international data on high school students performance. Any productivity study needs to account for all major inputs, and study time constitutes one of the most important inputs for education performance. A differences in main signaling stage across countries result in differences in this study time. Therefore, their estimates will be biased if they do not control for these differences.

This paper does not deal with the interesting issue of welfare comparison across the equilibria. Which equilibrium would be more desirable to have in our society? Our model is limited in that respect due to several simplifying assumptions. For example, we assume that studying either in high school or college does not improve human capital at all. In order to give a meaningful answer to this welfare comparison issue one needs to tackle related empirical questions. How much does education improve workers' productivity? Between high school and college education, which is more important? Between 20% of well educated college graduates and 90% of high school graduates, which is better for society? How much is the other cost of studying than students' time? These are challenging, but interesting questions for future research.

6 APPENDIX: Proofs of Lemma 1, Lemma 2, and Lemma 3

6.1 Proof of Lemma 1

Proof. Step 1: For all (s^*, p_c^*) such that a positive measure of workers attend college s^* and perform p_c^* in equilibrium, there exists $w(s^*, p_c^*) \in \mathbb{R}_+$ such that $w(s^*, p_c^*) \equiv w_1(s^*, p_c^*) = w_2(s^*, p_c^*)$.

Proof) Suppose to the contrary that $w_i(s^*, p_c^*) > w_{-i}(s^*, p_c^*)$ for some $i \in \{1, 2\}$. Alternative wage schedule $w_i' : w_i'(s^*, p_c^*) = (w_i(s^*, p_c^*) + w_{-i}(s^*, p_c^*))/2$ and $w_i' = w_i$ elsewhere, is a profitable deviation for firm i.

Step 2: For all (s^*, p_c^*) such that a positive measure of workers attend college s^* and perform p_c^* in equilibrium, it holds that $w(s^*, p_c^*) = E(\theta|s^*, p_c^*) + \alpha E(\theta|s^*)$.

Proof) Suppose that $w(s^*, p_c^*) > E(\theta|s^*, p_c^*) + \alpha E(\theta|s^*)$. An alternative wage schedule $\hat{w}_1$ for firm 1, where $\hat{w}_1(s^*, p_c^*) = E(\theta|s^*, p_c^*) + \alpha E(\theta|s^*)$ and $\hat{w}_1 = w$ elsewhere, constitutes a profitable deviation for firm 1. Suppose that $w(s^*, p_c^*) < E(\theta|s^*, p_c^*) + \alpha E(\theta|s^*)$. There exists $\varepsilon > 0$ such that an a alternative wage schedule $\hat{w}_1$ for the firm 1, where $\hat{w}_1(s^*, p_c^*) = w(s^*, p_c^*) + \varepsilon$ and $\hat{w}_1 = w$ elsewhere, constitutes a profitable deviation for firm 1. ■

6.2 Proof of Lemma 2

Proof. Since v is strictly concave, we obtain $\frac{\partial}{\partial \theta}\left(\frac{\partial \hat{U}/\partial p_c}{\partial \hat{U}/\partial w}\right) = -\frac{p_c}{\theta^3} v''\left(\frac{p_c}{\theta}\right) > 0$. ■

6.3 Proof of Lemma 3

Lemma A 1 *(i)* $C_A(x;\alpha)$ *is continuous and increasing in* x *for all* $\alpha > 0$*, and* $C_A(x;\alpha)$ *is a continuous and increasing in* α *for all* $x > 0$*.*

(ii) $\lim_{x\searrow 0} C_A(x;\alpha) = 0$ *for all* $\alpha > 0$*, and* $\lim_{\alpha\searrow 0} C_A(x;\alpha) = 0$ *for all* $x > 0$*.*

Proof. *(i)* C_A is uniquely determined by condition(5). C_A is continuous in x because $\tilde{\gamma}^i(x, C_A)$, defined in condition (4), and F are continuous. In order to satisfy condition (5) C_A has to increase when x increases, because $\tilde{\gamma}^i(x, C_A)$ is increasing in x and decreasing in C_A from condition (4). Therefore, C_A is an increasing in x. Similar arguments show that for all $x > 0$ $C_A(x;\alpha)$ is a continuous and increasing function in α.

(ii) Suppose that $\lim_{x\searrow 0} C_A(x;\alpha) \neq 0$ for some $\alpha > 0$. It follows from condition (4) that $\lim_{x\searrow 0} \tilde{\gamma}^i(x, C_A) = 0$ for $i = H, M, L$, and therefore $\lim_{x\searrow 0} \sum_{i=H,M,L} F\left(\tilde{\gamma}^i(x, C_A)\right) = 0$. This is a contradiction to condition (5). Similar arguments show that $\lim_{\alpha\searrow 0} C_A(x;\alpha) = 0$ for all $x > 0$. ■

Lemma A 2 *For all* $\alpha > 0$*, both* $\frac{\tilde{\gamma}^H(x;\alpha)}{\tilde{\gamma}^M(x;\alpha)}$ *and* $\frac{\tilde{\gamma}^M(x;\alpha)}{\tilde{\gamma}^L(x;\alpha)}$ *are increasing in* x*.*

Proof. It follows from condition (4) that

$$\frac{\tilde{\gamma}^H(x;\alpha)}{\tilde{\gamma}^M(x;\alpha)} = \frac{v(0) - v\left(C_A/\theta^M\right)}{v(0) - v\left(C_A/\theta^H\right)}.$$

Therefore, in order to show that $\frac{\tilde{\gamma}^H(x;\alpha)}{\tilde{\gamma}^M(x;\alpha)}$ is increasing in x, it suffices to show that $\frac{v(0)-v\left(C_A/\theta^M\right)}{v(0)-v\left(C_A/\theta^H\right)}$ is increasing in C_A, because C_A is an increasing function in x due to the part *(i)* of Lemma A1.

$$\begin{aligned}
\frac{\partial}{\partial C_A}\frac{v(0) - v\left(C_A/\theta^M\right)}{v(0) - v\left(C_A/\theta^H\right)} &= \left\{\begin{array}{c} \left(v(0) - v\left(C_A/\theta^M\right)\right)\cdot v'\left(C_A/\theta^H\right)\cdot\frac{1}{\theta^H} \\ -\left(v(0) - v\left(C_A/\theta^H\right)\right)\cdot v'\left(C_A/\theta^M\right)\cdot\frac{1}{\theta^M} \end{array}\right\} / \left(v(0) - v\left(C_A/\theta^H\right)\right)^2 \\
&> 0
\end{aligned}$$

The last line comes because $v(0) - v\left(C_A/\theta^H\right) > v(0) - v\left(C_A/\theta^M\right) > 0$, $v'\left(C_A/\theta^M\right) < v'\left(C_A/\theta^H\right) < 0$

($\because v$ is concave) and $1/\theta^M > 1/\theta^H > 0$. Similar arguments show that $\frac{\tilde{\gamma}^M(x;\alpha)}{\tilde{\gamma}^L(x;\alpha)}$ are increasing in x. ■

Lemma A 3 *(i) For all $\alpha > 0$, $\lim_{x \searrow 0} \frac{\tilde{\gamma}^H(x;\alpha)}{\tilde{\gamma}^M(x;\alpha)} = \frac{\theta^H}{\theta^M}$ and $\lim_{x \searrow 0} \frac{\tilde{\gamma}^M(x;\alpha)}{\tilde{\gamma}^L(x;\alpha)} = \frac{\theta^M}{\theta^L}$.*
(ii) For all $x > 0$, $\lim_{\alpha \searrow 0} \frac{\tilde{\gamma}^H(x;\alpha)}{\tilde{\gamma}^M(x;\alpha)} = \frac{\theta^H}{\theta^M}$ and $\lim_{\alpha \searrow 0} \frac{\tilde{\gamma}^M(x;\alpha)}{\tilde{\gamma}^L(x;\alpha)} = \frac{\theta^M}{\theta^L}$.

Proof. (i)

$$\lim_{x \searrow 0} \frac{\tilde{\gamma}^H}{\tilde{\gamma}^M} = \lim_{C_A \searrow 0} \frac{v(0) - v\left(C_A/\theta^M\right)}{v(0) - v\left(C_A/\theta^H\right)} = \frac{\theta^H}{\theta^M}$$

The first equation comes from condition (4) and Lemma A1. The second equation comes from the l'Hôpital's rule. Similar arguments show $\lim_{x \searrow 0} \tilde{\gamma}^M/\tilde{\gamma}^L = \theta^M/\theta^L$. Part (ii) can be proven in the same way. ■

Lemma A 4 *For all $\alpha > 0$, $\lim_{x \searrow 0} \tilde{\gamma}^L(x;\alpha) > \underline{\gamma}$ and $\lim_{x \searrow 0} \tilde{\gamma}^H(x;\alpha) < \bar{\gamma}$.*

Proof. Suppose that $\lim_{x \searrow 0} \tilde{\gamma}^L(x) \leq \underline{\gamma}$. It implies that no low ability workers exist in college A for x near 0. Since the capacity of college A is 1.5, college A has to have at least half unit measure of medium ability workers. Thus, $\lim_{x \searrow 0} \tilde{\gamma}^M$ has to be weakly greater than γ_m where $F(\gamma_m) = 0.5$, and we obtain

$$\lim_{x \searrow 0} \frac{\tilde{\gamma}^M(x;\alpha)}{\tilde{\gamma}^L(x;\alpha)} \geq \frac{\gamma_m}{\underline{\gamma}}.$$

However, using the part (i) of Lemma A3 and Assumption 1 we obtain the following contradicting result.

$$\lim_{x \searrow 0} \frac{\tilde{\gamma}^M(x;\alpha)}{\tilde{\gamma}^L(x;\alpha)} = \frac{\theta^M}{\theta^L} < \frac{\gamma_m}{\underline{\gamma}}.$$

Similar arguments show that $\lim_{x \searrow 0} \tilde{\gamma}^H(x) < \bar{\gamma}$. ■

Lemma A 5 *There exists $\gamma_0^i \in \left(\underline{\gamma}, \bar{\gamma}\right)$ such that (a) $\lim_{x \searrow 0} \tilde{\gamma}^i(x;\alpha) = \gamma_0^i$ for all $\alpha \geq 0$ (b) $\lim_{\alpha \searrow 0} \tilde{\gamma}^i(x;\alpha) = \gamma_0^i$ for all $x > 0$ $(i = H, M, L.)$*

Proof. The part (i) of Lemma A3 implies that $\lim_{x \searrow 0} \tilde{\gamma}^L(x;\alpha) < \lim_{x \searrow 0} \tilde{\gamma}^M(x;\alpha) < \lim_{x \searrow 0} \tilde{\gamma}^H(x;\alpha)$ for $\alpha > 0$. Therefore, Lemma A4 implies that $\underline{\gamma} < \gamma_0^i \equiv \lim_{x \searrow 0} \tilde{\gamma}^i(x;\alpha) < \bar{\gamma}$ for $\alpha > 0$ $(i = H, M, L.)$ Lemma

A3 also implies that the limiting ability distribution is identical whether x converges to 0 with α fixed or α converges to 0 with x fixed. Therefore $\lim_{x\searrow 0} \tilde{\gamma}^i(x;\alpha)$ with $\alpha > 0$ fixed is equal to $\lim_{\alpha\searrow 0} \tilde{\gamma}^i(x;\alpha)$ with $x > 0$ fixed $(i = H, M, L.)$ ■

6.3.1 Proof for part (*i*) of Lemma 3

Proof. Lemma A5 encompasses the part (*i*) of Lemma 3. ■

6.3.2 Proof for part (*ii*) of Lemma 3

Proof. C_A is continuous in x according to Lemma A1. We thus obtain from condition (4) that $\tilde{\gamma}^i(x;\alpha)$ is continuous in x $(i = H, M, L.)$ It follows from condition (6) that ψ is continuous in x. Lemma A2 implies that there will be weakly more higher ability workers in college A relative to lower ability workers as x increases. Therefore, ψ is weakly increasing in x. ■

6.3.3 Proof for part (*iii*) of Lemma 3

Proof. Part (*iii*) of Lemma 3 directly follows from Lemma A5. ■

7 References

Austen-Smith, D. (2002), "Peer Pressure and Job Market Signaling." Center for Mathematical Studies in Economics and Management (Northwestern University) Discussion Paper No. 1352.

BBC. (2000), "Chinese Police Probe Exam Scam." July 11; Online at http://news.bbc.co.uk/1/hi/world/asia-pacific/829015.stm.

Banks, J. S. and Sobel, J. (1987), "Equilibrium Selection in Signaling Games." *Econometrica*, 55(3), pp. 647-61.

Cho, I. and Kreps, D. M. (1987), "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics*, 102(2), pp. 179-221.

Cho, I. and Sobel, J. (1990), "Strategic Stability and Uniqueness in Signaling Games." *Journal of Economic Theory*, 50(2), pp. 381-413.

Elliott, D. (1999), "Learning to Think." *Newsweek*, September 6.

Engers, M. (1987), "Signalling with Many Signals." *Econometrica*, 55(3), pp. 663-74.

Gluck, C. (2001), "S Korean Students Face Do-or-Die Exam." *BBC*, November 6; Online at http://news.bbc.co.uk/1/hi/world/asia-pacific/1641670.stm.

Gold, T., Guthrie, D., Wank, D., and Granovetter, M. (2002), *Social Connections in China: Institutions, Culture, and the Changing Nature of Guanxi*, Cambridge University Press.

Hanushek, E. A. and Luque, J. A. (2003), "Efficiency and Equity in Schools around the World." *Economics of Education Review*, 20(4).

Heyneman, S. P., and Loxley, W. (1983), "The Effect of Primary School Quality on Academic Achievement across Twenty-Nine High and Low Income Countries." *American Journal of Sociology*, 88, pp. 1162-94.

Japanese Statistics Bureau (2002), *The 2001 Survey on Time Use and Leisure Activities.*, Online at http://www.stat.go.jp/english/data/shakai/index.htm.

Kim, K., Kim, S., and Nahm, K. (1999), "Seoul National University Students Life Pattern Survey." Center for Campus Life and Culture, Seoul National University.

Lafayette, B. and Mente, D. (1994), *Japanese Etiquette & Ethics in Business (6th edition)*, McGraw-Hill/Contemporary Books, 1994.

Lahmers, A. G. and Zulauf, C. (2000), ("Factors Associated with Academic Time Use and Academic Performance of College Students: A Recursive Approach." *Journal of College Student Development* 41(5), pp. 544-556.

Lazear, E. P. (2001), "Educational Production." *Quarterly Journal of Economics*, 116(3), pp. 777-803.

Light, K. W. and Wais, M. (2000), "Stanford 2000 Student Time-Use Study" Online at http://www.stanford.edu/dept/DOS/studentsurvey/.

OECD. (2001), "Knowledge and Skills for Life: First Results from Pisa 2000."

Park, S. (2003), "Not everyone is welcome to apply." *JoongAng Daily*, December 3; Online at http://joongangdaily.joins.com/200312/03/200312030152038679900092109211.html.

Quinzii, M. and Rochet, J. (1985), "Multidimensional Signalling." *Journal of Mathematical Economics*, 14(3), pp. 261-84.

Riley, J. G. (1975), "Competitive Signalling." *Journal of Economic Theory*, 10(2), pp. 174-86.

Spence, A. M. (1973), "Job Market Signaling." *Quarterly Journal of Economics*, 87(3), pp. 355-74.

Walberg, H. J. (2001), "Do American Students Study Enough?" *Hoover Institution Weekly Essays*, July 23; Online at http://www-hoover.stanford.edu/pubaffairs/we/current/walberg_0701.html.

Woessmann, L. (2000), "Schooling Resources, Educational Institutions, and Student Performance: The International Evidence." *Kiel Working Paper* No. 983, Kiel Institute of World Economics.