Imperfect Information, Self-Selection and the Market for Higher Education

Tal Regev*
Department of Economics, MIT
Job Market Paper

November 22, 2005

Abstract

This paper explores how the steady trends in increased tuition costs, college enrollment and returns to education might be related to the quality of college graduates. The model shows that the signaling role of education might be an important, yet largely neglected ingredient in these recent changes. In a special signaling model, workers face the same costs, but can expect different returns from college. Allocation of ability into skill is determined by the equilibrium skill premium. Incorporating a production of higher education, the properties of the college market equilibrium are discussed. A skill biased technical change initially decreases self-selection into college, but the general equilibrium effect can overturn the initial decline, since increased enrollment and rising tuition costs increase selection. Higher initial human capital has an external effect on subsequent investment. All agents increase their schooling investment, and the higher equilibrium tuition costs increase self-selection and the college premium.

*Email talir@mit.edu. I am grateful to my advisors Daron Acemoglu, Ricardo Caballero and Ivan Werning for their helpful suggestions and support. I also thank David Autor, Karna Basu, Mike Golosov, Guido Lorenzoni, Veronica Rappoport and seminar participants at MIT for their comments.

1 Introduction

Since mid-century there has been a steady increase in the returns to college, which was interrupted between 1971 to 1979, and then became quite steep since the late 70's (see figure[1]). This increase reflects a real increase in college graduates wages and at times a real decline in high-school graduates wages[1]. The number of college graduates followed this trend closely, rising from 8% of 24 years and older in 1960 to 25% in 1997 (see figure[2])[2]. The third phenomenon is an increase in college costs, which have increased more than two fold during the past thirty years (see figure[3]). This paper explores how these changes in the college market are related to the quality and self-selection of college graduates. While issues of selection of heterogeneous workers across education levels and occupations has received much attention, these studies typically assume that workers sort themselves into school according to an exogenous distribution of costs and returns. The growing body of evidence suggesting a large signal component to education has been largely ignored[3]. Nevertheless, the signal role of college education might be an important ingredient in the recent changes in the demand for college education and the structure of the college market. This paper constructs a signaling model of the demand for skills, skill premium and college market equilibrium..

Perhaps one reason the original Spence (1973) model was not fully applied to the college market is the assumption that workers' ability is negatively correlated with their costs of attending college. It seems unlikely that low ability workers face substantially higher costs than able workers. On the contrary, if a large part of the cost of going to college are forgone earnings, the correlation between ability and costs could even be positive[4]. This paper departs from the standard signalling framework by assuming workers face the same costs but can expect different returns from college, due to the fact that an able worker is more likely to be perceived as such[5]. Allowing some information on a worker's ability reintroduces a form of single crossing property, which is responsible for the resulting self-selection. By keeping information limited, firms still need to infer a worker's ability using the composition of his skill group, thus preserving the value of education as a signal.

We explore the properties of this special signaling equilibrium, and show that a separating

[1] According to Juhn, Murphy and pierce (1993) the college to high school wage ratio was 1.45 in 1979, and soared to 1.9 in 1986. Katz and Murphy (1992) show a decline in the wages of low skill workers between 1979 and 1987. See also Goldin and Katz (2000), Katz and Autor (1999), Autor, Katz and Kearny (2005) and Card (1999).

[2] US Bureu of the Census, 1998

[3] Tyler, Murnane and Willett (2000) find that a General Educational Development signal increases wages by 10-19% net of human capital effects. Lang and Kropps (1986) provide more direct evidence on signaling as an equilibrium phenomenon, and show compulsory schooling laws affect attendance decisions even for non-marginal agents. In the same spirit Bedard (2001) shows that high school dropout rates increase when the pool of high school graduates deteriorates.

[4] A point stressed by Griliches (1977) for instance.

[5] See also Arrow (1973b) for educational signaling and Weiss (1986) and Hvide (2003) for a setup similar to mine.

equilibrium does not exist. The fact that a Riley (1979) critique does not apply provides some justification for focusing on the totally mixed equilibrium[6]. This equilibrium also seems plausible from an empirical viewpoint, and is more interesting because the distribution of abilities is non-trivial. We show than an equilibrium with positive ability selection into skill arises when it is socially inefficient for the low ability worker to invest in schooling. We proceed to show that self selection increases when the cost of college is higher or when productivity of college graduates is lower. Finally we show that total investment in education may actually increase with college costs. Behind this result are strategic externalities. Lower net returns to college makes schooling less attractive for all workers. When low ability workers shift away from college, the signaling value of college increases and the value of remaining unskilled simultaneously declines. If the relative increase in the value of skill dominates the original increase in tuition costs, enrollment rates can be higher. In such a case the demand for college will be upward sloping.

The model also predicts human capital externalities. All workers invest more in their human capital when the average human capital is initially high. The equilibrium composition of each skill level is given such that workers are indifferent between the skill choices, and it is invariant with the initial level of ability. When there are more able workers in the population, all workers must increase their investment in education to keep the same equilibrium proportion intact. Groups with higher initial human capital will invest more in education, a fact which has a dynamic implication of increasing divergence between groups. This result is reminiscent of the statistical discrimination literature[7]. However, the multiple equilibria and coordination failure assumption driving these results is not present here. Closer in spirit is Acemoglu (1996) where a different mechanism results in increasing returns to human capital[8].

To close the college market, a college supply function is introduced. We assume college production uses some scientists who are in limited supply, and whose wage is determined in equilibrium. Combining this with the demand for college we solve for the equilibrium in the college market. We use this framework to describe the recent changes in the college market with the augmented selection prediction. An increase in initial human capital has no first order effect on selection, but through increased investment the price of college increases, resulting in higher self-selection, lower wages for unskilled workers and a higher college premium. More complex are the effects of a skill biased technical change (SBTC) which is widely believed to have taken place during this period[9]. The direct increase in the skill premium following a SBTC is undermined by the lower quality of workers who now choose to become skilled. Within-skill (residual) inequality

[6] In the standard setup, the only equilibrium surviving refinements, such as the Cho Kreps (1987) critirion, is the Riley equilibrium which is the best separating equilibrium.

[7] Arrow (1973a), Phelps(1972), Coate and Loury (1993)

[8] In Acemoglu (1996) the increasing returns stem from ex-ante investment and costly search.

[9] See Acemoglu (2002) for a comprehensive study. See also Autor, Katz and Krueger (1998) and Autor, Levy and Murnane (2003) on the skill content of computerization. Krusell and al. (2000) use capital-skill complementarities to estimate the skill biased change. Berman Bound and Griliches (1994) bring independent evidence from the manufacturing sector pointing to an increase in SBTC in the 70's.

increases, while the college premium declines, as has been the case throughout the 70's. When enrollment and tuition increase, the general equilibrium effect can overturn the initial decline in selection resulting in an increase of the skill premium. Such higher self selection can possibly account not only for the wage increase of high skilled workers, but also for the reduced wages of the uneducated, a fact which otherwise remains a puzzle[10].

Almost all studies find a positive selection bias[11]. More controversial is the evidence regarding the dynamic change in selection over time. Cameron and Heckman (1998) report a decline in the quality of college graduates. Juhn Kim and Vella (2005) suggest a smaller decline in the quality of younger more educated cohorts. Card and Lemieux (2001) find new cohorts to have higher returns, but do not interpret this as an increase in ability component. Murnane, Willett and Levy (1995) find an increase in the ability composition of educated workers[12]. However, these studies are primarily interested in the returns to skill and hence estimate only the composition of the skilled labor force. Nevertheless, the model predicts changes in the relative composition of the skilled and unskilled pools. The strongest evidence in favor of increased ability selection into skill is given in Steinberger (2005). Using direct new data on test scores in 1979 and 1999 he reports a 4% rise in ability for male graduates with a simultaneous decline in the ability of high school graduates. A direct test of the mechanisms leading to such increased selection is not available yet.

Treatments of the whole college market are few. Hoxby (1997) investigates the supply side of college and its increased differentiation and competition. Rothschild and White (1995) present a pricing model of college. There is ongoing interest in the demand side and in particular on the effect of tuition subsidies (Feldstein (1995)). Hendel and al. (2005) look at the effect of subsidies on inequality in a signaling equilibrium with credit constraints. There has not been an attempt to look at the college market equilibrium within a signaling perspective.

The rest of the paper proceeds as follows. Section 1 presents the model and solves for the signaling equilibria. Section 2 analyses the mixed strategy equilibria. Section 3 discusses self selection, the skill premium, the investment decision of workers and welfare. Section 4 extends to college production and the equilibrium in the college market. Section 5 concludes.

---

[10] See Autor Katz and Kearny (2005) for an estimate of the declining prices for low skill workers. Autor, Levy and Murnane (2003) suggest a possible explanations for the polarization of the labor market can be computerization. See Acemoglu (1999) for an alternative explanation of how a SBTC changed the composition of jobs, reducing the wages of low skill workers.

[11] These conclusions aggregate over findings of large selection bias (Blackburn (1995)), and a small negative bias (Angrist and Krueger (1991)). The measure of selection is usually derived from a comparison of the OLS estimate with the unbiased IV estimate of the returns to skill, and the size of the estimated bias depends on institutional change used as the instrument. A different idententification is given in Ashenfelter and Rouse (1998) who use a sample of identical twins to estimate asmall upward ability bias. See Card (2002) for a complete survey.

[12] Cameron and Heckman (1998) find that the location of average ability of graduates in the baseline distribution has steadily declined from .92 to .85 during the course of fifty years (for the cohorts born in 1916 to those born in 1963). Card and Lemieux (2001) interpret their findings as arrising from complementarities between cohorts. Murnane Willet and Levy (1995) use direct test scores measures to control for the ability bias.

2 Model

Workers of heterogenous abilities choose whether to acquire costly education which is productivity augmenting. Firms observe the worker's schooling choice and an additional proxy on a worker's ability. Wages are determined in a competitive labor market. We assume for simplicity that workers of various abilities and skill levels are perfect substitutes. This simplifying assumption makes equilibrium wages depend in effect only the quality of workers with the same observed skill level, and leaves out the standard quantity effects[13].

2.1 Setup

The economy consists of a continuum of mass M of firms, and a unit mass of risk neutral workers with heterogenous abilities. We identify each type of worker $i \in \{h, l\}$ by his ability $a_i \in \Theta = \{a_h, a_l\}$. A worker's type is private information. The prior distribution of types in the population is common knowledge and is given by $(p(a_h), p(a_l)) = (f, 1 - f)$.

A worker of type i chooses his skill level $e \in E = \{L, H\}$. A strategy for worker i is a probability distribution over the set of actions, $\sigma_i : E \longrightarrow [0, 1]$.

All firms observe each worker's skill choice e and an additional noisy signal $s \in S = \{h, l\}$ on the worker's true ability. Let the likelihood of a worker i emitting a signal s be given by the distribution $p_i : S \longrightarrow [0, 1]$ which is also common knowledge. A Firm's strategy is a wage offer $w_e(s)$ where $w : E \times S \longrightarrow R^+$.

A firm that hires l_e^i workers of type i and skill level e produces output via a linear technology $y = \sum_i a_i(\lambda l_H^i + l_L^i)$ where $\lambda \geq 1$. That is, each worker's marginal productivity is his ability, enhanced by a multiplicative premium of λ if the worker is skilled[14].

Each firm's payoff from hiring a worker with observable (e, s) is given by the quadratic loss function $\pi(e, s, w) = (w_e(s) - \lambda_I E\left(a|e, s\right))^2$ where $\lambda_I = \lambda$ if $e = H$ and 1 otherwise, which is the standard shortcut way to replicate a competitive labor market outcome. Worker's i payoff from any pure action e is given by $u_i(e, s, w) = w_e(s) - C(e)$, where we normalize $C \equiv C(H) > C(L) = 0$. Note workers of different types get the same utility. Their payoffs only differ in expectation, $Eu_i(e, s, w) = \sum_s p_i(s) w_e(s) - C(e)$.

We assume the signal is informative in the sense that the likelihood of a good signal for the high ability worker is larger than the likelihood of a good signal for a low ability worker,

Assumption A1 (*Monotone Likelihood*): $\frac{p(s|a_h)}{p(s|a_l)}$ increases with s.

The assumption that signals depend only on the worker's ability and not on his skill choice is

[13] A natural extention would be to allow for some complementarity between skill levels, and incorporate also the quantity effects. See Moro and Norman (2004) for a general equilibrium model of missing infromation and production complementarities.

[14] We could allow skill to affect worker's productivities differentially, i.e., have $\lambda_l \lesseqgtr \lambda_h$ without any substantial changes to the results.

only made for simplicity[15]. Throughout we use lower case $\{h, l\}$ to denote the signal on abilities, while skill levels are denoted by upper case $\{H, L\}$.Whenever there is no ambiguity we will use the shortcut notation $(\sigma_i, 1-\sigma_i) \equiv (\sigma_i(H), \sigma_i(L))$.

2.2 Equilibrium Definition

Let $\mu(\cdot|e, s)$ denote the posterior distribution over types $\{a_h, a_l\}$ after observing (e, s).

Definition 1 *A Perfect Bayesian Equilibrium of this game is a tuple* $\{\sigma^*, w_e^*(s), \mu^*(\cdot|e, s)\}$ *of workers' strategies and firms' wage offers and beliefs such that:*

1. *Workers maximize their expected payoffs:*
   $\forall i, \sigma_i^*(\cdot) = \arg\max \sum_s p(s|a_i)\left[\sigma_i(H)\left(w_H^*(s) - C\right) + (1-\sigma_i(H))w_L^*(s)\right]$
2. *Firms pay the workers their expected productivity:*
   $w_e^*(s) = \lambda_I \sum_{a_i} \mu^*(a_i|e, s)a_i$
3. *Posterior beliefs are Bayesian wherever possible:*
   $\mu^*(a_i|e, s) = p(a_i)\sigma_i^*(e)p(s|a_i) / \left(\sum_{a_{i'}} p(a_{i'})\sigma_i^*(e)p(s|a_{i'})\right)$
   if $\sum_{a_{i'}} p(a_{i'})\sigma_i^*(e)p(s|a_{i'}) > 0$, *and any probability distribution over* $\{a_h, a_l\}$ *otherwise.*

Denote the expected wage a worker of type a_i gets if he chooses skill level e by $Ew_e(a_i) \equiv \sum_s p(s|a_i)w_e^*(s)$. A worker's expected wage from a skill choice e depends on his own type through the probability term $p(s|a_i)$, and on the equilibrium composition of his skill group through the equilibrium wage term $w_e^*(s)$. This is the feature of the model which allows worker's returns to increase with ability, reintroducing a type of single crossing property, while keeping the information externality intact. The inferred ability of a worker still depends on the equilibrium composition. If in equilibrium the posterior probability of finding a high ability worker is higher in the skilled group than in the unskilled group, then acquiring skill has an additional signaling value.

2.3 Types of Equilibria

A worker of ability a_i compares his expected payoff when he acquires skill, $Ew_H(a_i) - C$, and his expected payoff when he does not, $Ew_L(a_i)$, and chooses the education level with the higher returns given equilibrium play of all other agents. If he is indifferent between the choices then any mixed strategy is (weakly) optimal. Each type's strategy represents the fraction of the population of that type that plays the pure strategy skill choice. Because we are interested in a

[15] A realistic modification will assume that the signal is more precise for high skilled workers, that is: $p_{h|H} > p_{h|L}$ and $p_{l|H} < p_{l|L}$.This will result in different within skill variance.

non-trivial composition of skills, we are mainly interested in the interior (fully mixed strategy) solution. An additional theoretical justification for studying the interior equilibrium is that there is no separating equilibrium, as we show below.

We briefly characterize the full equilibrium possibilities. We begin with two Lemmas:

Lemma 1 *$w_e^*(s)$ increases with s.*

Proof. $w_e^*(s) = \lambda_I \sum_{a_i} \mu^*(a_i|e,s)a_i$ is an increasing function of beliefs and by the monotone likely hood assumption ($A1$), beliefs are an increasing function of the signal, s, i.e., for all equilibrium beliefs formed by Bayes rule, $\mu^*(a_h|H,h) \geqslant \mu^*(a_h|H,l)$ and $\mu^*(a_h|L,h) \geqslant \mu^*(a_h|L,l)$ with strict inequality for interior beliefs $\mu^*(a_h|\cdot) \notin \{0,1\}$. ■

Lemma 2 *$Ew_e(a_h) \geqslant Ew_e(a_l)$*

Proof. Follows from previous lemma by the monotone likelihood assumption. ■

We use these to prove:

Proposition 1 *There is no equilibrium in which the high-ability worker reveals himself.*

Proof. Assume to the contrary that a_h reveals himself in skill level e. By Bayes rule $\mu^*(a_h|e,h) = \mu^*(a_h|e,l) = 1$ so that $w_e(h) = w_e(l) = w_e$. By the previous Lemma, in the other sector $\widetilde{e}$ we have $Ew_{\widetilde{e}}(a_l) \leqslant Ew_{\widetilde{e}}(a_h) \leqslant Ew_e(a_h) = Ew_e(a_l) = w_e$ where the second inequality follows from a_h's choice and the equality from the beliefs being $\mu^*(a_h|e,\cdot) = 1$. This contradicts $Ew_{\widetilde{e}}(a_l) > Ew_e(a_l)$. ■

There can be no separating or semi-separating equilibrium in which the high type reveals himself. The proof provides the intuition: if there was an equilibrium in which the high-type reveals himself in e then Bayes rule would dictate that firms believe a worker is high ability when they observe skill choice e, regardless of the ability-signal. But if the ability-signal has no power, there is nothing keeping the low ability type from imitating the high-ability type. In fact, he will weakly prefer to do so, since he always does worse than the high ability type by choosing $\widetilde{e}$ where the ability-signal has power[16].

For the sake of completeness we now characterize the full set of equilibria. In what follows let $A \equiv \frac{C - a_l(\lambda-1)}{a_h - a_l}$, $a \equiv \frac{p_l}{p_h}$, $b \equiv \frac{1-p_l}{1-p_h}$.

Proposition 2 *Characterization of equilibria in terms of workers' strategies.*

(i) (Fully mixed strategy): An equilibrium with $(\sigma_l, \sigma_h) \in (0,1)^2$ exists only if $\frac{\lambda-(A+1)}{\lambda-A}\frac{A}{A+1} > \frac{4ab}{(a+b)^2}$.

(ii) (Pooling on H): An equilibrium with $(\sigma_l, \sigma_h) = (1,1)$ exists iff $\frac{p_h p_l}{p_h+p_l} + \frac{(1-p_h)(1-p_l)}{1-p_h+1-p_l} > \frac{A}{\lambda}$

[16] This result might be viewed as a weakness because it implies a discontinuity of the solution in the neighborhood of perfect information. We address this issue in the appendix.

(iii) (Pooling on L): An equilibrium with $(\sigma_l, \sigma_h) = (0,0)$ *exists* iff $-\left(\frac{p_l p_h}{p_h + p_l} + \frac{(1-p_l)(1-p_h)}{(1-p_h)+(1-p_l)}\right) < A$.

(iv) (a_l reveals himself in L): An equilibrium with $\sigma_l \in (0,1), \sigma_h = 1$ *exist* iff *there is a solution to* $\frac{p_l p_h}{p_h + \sigma_l p_l} + \frac{(1-p_l)(1-p_h)}{(1-p_h)+\sigma_l(1-p_l)} = \frac{A}{\lambda}$

(v) (a_l reveals himself in H):: An equilibrium with $\sigma_l \in (0,1), \sigma_h = 0$ *exist* iff *there's a solution to* $-\frac{p_l p_h}{p_h + (1-\sigma_l) p_l} - \frac{(1-p_l)(1-p_h)}{(1-p_h)+(1-\sigma_l)(1-p_l)} = A$ *(which can happen only If* $A < 0$*).*

And there is no other equilibrium

Proof. See appendix. ■

There are basically three types of equilibria. The fully mixed strategy, two pooling equilibria, and two equilibria where the high ability worker plays a pure strategy and the low ability worker mixes (and hence reveals himself). These equilibria exist in different regions (not mutually exclusive) of the parameter space. The four dimensional parameter space consists of the information probabilities p_h and p_l, the skill technology term λ, and the term A, which we interpret below as the social cost of having low types invest in school.

Both types can pool on skill if the social cost of low types investing in skill is small enough. On the other hand, pooling on L exists if there's a cost associated with low-type investing in skill, or if the gains from investing are small enough. However, these two pooling equilibria are not very robust to refinements that use forward-induction type arguments.

To be concrete, the pooling equilibria fail the divinity criterion (Banks and Sobel (1987))[17]. According to the divinity criterion, we can eliminate an equilibrium if we can show that there are beliefs regarding off-the-equilibrium-path skill choice for which only one type of worker would like to deviate. To adapt their criterion to this setting with minimal notations we define $D(a_i, e)$ to be the set of beliefs which makes type a_i strictly prefer deviating to e over his equilibrium strategy σ^*. Define $D^0(a_i, e)$ the set of beliefs for which type a_i is exactly indifferent.

Definition 2 *A type* a_i *is deleted for strategy* e *under the divinity criterion if there is another type* a_j *s.t.* $D(a_i, e) \cup D^0(a_i, e) \subset D(a_j, e)$

In other words, if type a_i is willing to deviate for a strictly smaller set of beliefs, then firms should believe that type a_j is the one deviating. The pooling equilibrium is thus destroyed.

Proposition 3 *The pooling equilibria do not survive the divinity criterion*

Proof. See appendix. ■

[17] Note that a pooling equilibrium does not fail the slightly weaker Cho-Kreps (1987) intuitive criterion. A pooling equilibrium fails the intuitive criterion if deviating is equilibrium dominated for the low type but the high type would prefer to deviate once the firm's beliefs assign probability zero to a low type deviating. In our case, a low type will like to deviate if the firm strongly believes a deviator is high ability. Hence the first part of the criterion is never satisfied.

These refinements, however, can only eliminate an equilibrium which has off-equilibrium belief assignments. They do not apply to the semi-separating and fully mixed equilibrium where beliefs are set by Bayes rule. Consider the semi-separating equilibrium. If $A > 0$ there's a social cost associated with the low ability worker getting skill, and the corresponding semi-separating equilibrium has all the high ability workers investing in skill. If it is socially efficient for the low ability worker to invest in skill ($A < 0$), then the corresponding semi-separating equilibrium has all the high ability workers not getting any skill.

We now turn to the fully mixed strategy.

3 Equilibrium Analysis

In the fully mixed strategy equilibrium each worker type is indifferent between the skill choices. To gain more intuition regarding the forces at work, we think about the solution in terms of the resulting quality in each skill level instead of the investment strategies σ_i. Explicitly writing the two equilibrium equations in terms of the quality variables we see there are at most two mixed-strategy equilibrium. We then discuss the properties of these equilibria

3.1 A Change of Variables: from Quantities to Qualities

Consider the following change of variables. Let ϕ denote the probability of finding a high-ability worker in the skilled pool, and similarly let ψ be the fraction of high ability workers in the unskilled group. That is

$$\begin{aligned} \phi &\equiv p(a = a_h|H) = \frac{f\sigma_h}{f\sigma_h + (1-f)\sigma_l} \\ \psi &\equiv p(a = a_h|L) = \frac{f(1-\sigma_h)}{f(1-\sigma_h) + (1-f)(1-\sigma_l)} \end{aligned} \tag{1}$$

These proportion of able workers in each skill group can be interpreted as the endogenous quality of the two groups. When firms make up their wage decision they take into account these variables, indicative of the group's composition. They then update this 'interim-prior' based on the additional individual ability-signal. We will see how this change of variables turns out to be quite useful, as the predictions regarding investment are limited, but the equilibrium forces are more easily interpreted through these quality variables.

3.2 Mixed Strategy Equilibrium

To solve for the equilibrium, we assume firms use workers' equilibrium mixing σ^* to form their correct beliefs $\mu^*(\cdot|e,s) = p(a_i)\sigma_i^*(e)p(s|a_i)/\left(\sum_{a_{i'}} p(a_{i'})\sigma_i^*(e)p(s|a_{i'})\right)$. Given these beliefs, we can construct expected wages for each skill level and ability-signal, $w_e^*(s) = \lambda_I \sum_{a_i} \mu^*(a_i|e,s)a_i$. Finally, workers take these wages as given when making their skill decision. In the fully mixed strategy equilibrium the expected returns from a worker's choices must be equalized. This boils down to two equilibrium equations, one for each type of worker i,

$$\sum_s p(s|a_i)w_H^*(s) - C = \sum_s p(s|a_i)w_L^*(s) \tag{2}$$

Plugging in the expressions for wages, these two equations solve for $\langle \sigma_h, \sigma_l \rangle$ provided they are between zero and one. If they are not, then a fully mixing equilibrium does not exist. Using our change of variables the equilibrium equations simplifies. After some algebra we get the equilibrium conditions,

$$\lambda \frac{\phi p_h}{\phi p_h + (1-\phi)p_l} = \frac{\psi p_h}{\psi p_h + (1-\psi)p_l} + \frac{(C - a_l(\lambda - 1))}{(a_h - a_l)} \tag{3}$$

$$\lambda \frac{\phi(1-p_h)}{\phi(1-p_h) + (1-\phi)(1-p_l)} = \frac{\psi(1-p_h)}{\psi(1-p_h) + (1-\psi)(1-p_l)} + \frac{(C - a_l(\lambda - 1))}{(a_h - a_l)} \tag{4}$$

Each equation correspond to an ability signal: the first equation equates the rewards for a high-signal ($s = h$) from getting skill or remaining unskilled and the second equation does the same for a low signal ($s = l$). In other words, the composition of the abilities in each skill-level ϕ and ψ must be such that the returns to an ability-signal are the same across skill levels. Both conditions impose a positive relationship between the quality of workers in the H and L sectors. This is because the investment decision of high ability workers has a positive external effect on the value of the skill group. To equate the returns across skill choices, ϕ and ψ must move together.

These expressions also disentangle the real value of skill from the informational value. Looking at each equation separately, the last term on the right, $A \equiv \frac{(C-a_l(\lambda-1))}{(a_h-a_l)}$ is the (normalized) net cost of having low ability workers invest in skill. It can be thought of as the social cost of having imperfect information . The other two terms of the form $F(q) = \frac{qp_1}{qp_1+(1-q)p_2}$ are the inference terms. The distribution of abilities must be such that the value added from information just compensates for the cost.

This representation also highlights the role of having some information available to firms. Suppose they had no information regarding workers' ability. In such a case, with $p_h = p_l$, the two conditions collapse to one and the equilibrium cannot be pinned down. Adding some information brings in a way to differentiate the returns of the two types, and reduces the number of equilibria substantially.

In fact, since the first equation is increasing and concave in $\langle \psi, \phi \rangle$ space, and the second equation is increasing and convex in $\langle \psi, \phi \rangle$ and since workers' mixed strategies uniquely define the Bayesian beliefs, this proves:

Proposition 4 *There are at most two mixed-strategy equilibria.*

4 Results: Self-Selection, Skill Premium, and Human Capital Investment

In this section we explore the implication of the equilibrium allocation and wages for our objects of interest. We show that self selection arrises if and only if it is inefficient for the low ability type to invest in skill. As for the comparative statics, a higher cost of education increases self selection, while a skill biased technical change reduces it. We show how self-selection translates directly into the skill premium. Next we look at investment and how it responds to price changes in the economy. We show that it increases in the initial level of human capital. We finish with a short discussion of welfare.

4.1 Self-Selection

Any solution $\langle \psi^*, \phi^* \rangle$ can be characterized by the degree to which high ability persons are more concentrated in the H sector. We define an index of self-selection as the difference between the proportion of high ability persons in the H sector and their proportion in the L sector.

Definition 3 *Let the measure of self selection be* $\phi - \psi$

We say that there is 'self-selection' or 'positive sorting' when the fraction of high ability workers is higher in the H sector than in the L sector. That is, if $\phi > \psi$ or alternatively $\sigma_h > \sigma_l$. A necessary and sufficient condition for self-selection to arise is the following condition, which we will now assume,

Assumption A2 : It is inefficient for the low ability type to invest in skill. $(C - a_l(\lambda - 1) < 0)$

This represents the net-normalized cost of having low ability workers pretend to be high ability. To make the model interesting we want to assume this cost is positive, that is, if information was perfect low ability workers would not find it rewarding to invest in skill. This assumption turns out to be crucial for self-selection to arise, as we now prove,

Proposition 5 *Self selection arrises* iff *it is inefficient for the lowest type to invest in skill (A2). That is,* $\phi > \psi \Leftrightarrow C > a_l(\lambda - 1)$

Proof. Assume $C > a_l(\lambda - 1) \Leftrightarrow \frac{C - a_l(\lambda-1)}{a_h - a_l} > 0$. Together with the first equilibrium equation [3] we have $\Longleftrightarrow \frac{\lambda\phi p_h}{\phi p_h + (1-\phi)p_l} > \frac{\psi p_h}{\psi p_h + (1-\psi)p_l}$. The two equilibrium equations [3] and [4] imply $\frac{\lambda\phi(1-\phi)}{\phi p_h + (1-\phi)p_l(1-\phi p_h - (1-\phi)p_l)} = \frac{\psi(1-\psi)}{\psi p_h + (1-\psi)p_l(1-\psi p_h - (1-\psi)p_l)}$. So that the last inequality holds iff $\frac{(1-\phi)}{(1-\phi p_h + (1-\phi)p_l)} < \frac{(1-\psi)}{(1-\psi p_h - (1-\psi)p_l)} \Longleftrightarrow \frac{\phi}{(1-\phi)} > \frac{\psi}{(1-\psi)} \Leftrightarrow \phi > \psi$ (or $\sigma_h > \sigma_l$). ■

Assumption A2 basically assures us that selection goes the right way. The assumption serves the same function as the standard single crossing assumption that costs decline with ability.

However, it is not its natural analogue. A first guess would have been that assumption 1 (monotone likelihood assumptions) is sufficient for self-selection since it assures us that the high ability worker gets higher returns from skill. However, recall that Assumption 1 implies that the high ability worker gets higher returns in any skill level (see Lemma 3).

We could also assume that it is efficient for the high ability type to invest in skill ($a_h < \lambda a_h - C$). However, we do not need this assumption. There can be a mixed equilibrium with some fraction of both types investing in skill even if it inefficient for both of them to invest.

For our next result on the comparative statics of self-selection, we need the following condition on the parameters:

Condition 1 $\frac{p_h p_l}{(1-p_h)(1-p_l)} < \frac{(fp_h+(1-f)p_l)^2}{((1-f)p_h-(1-f)p_l)^2}$.

Proposition 6 *self selection increases and the quality of unskilled workers declines* $((\phi - \psi) \uparrow$ *and* $\psi \downarrow)$ *with:*

(i) increased costs ($C \uparrow$)

(ii) decreased returns to skill ($\lambda \downarrow$)

(iii) decreased productivity of high-ability worker($a_h \downarrow$)

(iv) increased productivity of low-ability worker ($a_l \uparrow$)*, provided it is efficient for the high-ability type to invest in skill*

If Condition(2) holds[18].

Proof. See Appendix. ■

Anything which increases the costs of skill for the low ability worker, reduces his relative investment and increases self-selection. The empirical implications are straightforward. If we believe there has been an increase in the real skill premium we should expect the relative quality of high skill workers to decrease, as there will be relatively more low-ability workers trying to gain from the higher returns. The quality of low skill workers should decline in absolute terms. On the other hand, the on-going increase in education costs should increase self-selection, that is, the demand for college from quality applicants is increasing in tuition costs.

4.2 Skill Premium

Wages in this economy depend on the composition of the skill group through the information externality. In equilibrium wages are constructed as the expected productivity of an individual with a certain skill level and ability-signal. Workers benefit from an increase in the quality of their skill group regardless of the specific signal they turn out to emit:

Lemma 3 $w_e^*(s)$ *increases with* $p(a_h|e)$ *for all s.*

[18] This condition is only a convinient sufficient condition to proove this result. Rigorous simulation suggest the result holds under much narrower restrictions.

Proof. By construction. ■

All workers within a skill level benefit from its quality. Average wages increase with the quality of the group not simply as an averaging result. Rather, the higher average wage truly comes from increased wages for all workers in the skill group. The composition of the group actually affects prices and is not a phantom 'composition effect' we usually try to control for in our wage regressions.

Keeping this in mind, it is natural to define the skill premium as the difference between expected (average) wages of a skilled and unskilled worker, where expected wages are given by $Ew_e \equiv \sum_{a_i} p(a_i|e) \sum_s p(s|a_i) w_e^*(s). = \sum_{a_i} p(a_i|e) a_i$. The actual realization of the wage is just some noise around these means.

Definition 4 *The skill premium is defined as* $Ew_H - Ew_L$

And so we have,

Proposition 7 *The skill premium increases with selection.*

Proof. Writing out the expressions for expected wages we have $Ew_H = \lambda (\phi a_h + (1-\phi) a_l)$ and $Ew_L = \psi a_h + (1-\psi) a_l$ so that the skill premium

$$Ew_H - Ew_L = (\lambda\phi - \psi)(a_h - a_l) \tag{5}$$

increases with $\phi - \psi$. ■

Anything that increases selection, increases the skill premium. From proposition(6) we know what these are under some condition. However, we can prove a stronger unconditional result,

Proposition 8 (i) *An increase in the real returns to skill,* $\lambda \uparrow$, *or an increase in the ability gap,* $(a_h - a_l)\uparrow$ *directly increases the skill premium, but has an indirect dampening effect on the skill premium through decreased selection.*

(ii) *Higher investment costs increase the skill premium through the indirect increase in selection.*

Proof. See Appendix. ■

The first result suggests that a skill biased technical change would have created larger wage dispersion absent the endogenous selection adjustment. The relative shift of low ability workers into education reduces the relative quality of the skilled. Hence, a firm's willingness to pay for the higher productivity, λ, is diminished. The second result says that an increase in education costs improves the quality of college graduates, and thus indirectly increases their wages. Costs, which are uncorrelated with ability or wages, turn out to have an effect on wages. This results highlight an empirical implication of the model. We cannot control for selection when looking for the skill premium, because selection is part of what drives wages.

4.3 Investment in Human Capital

The comparative statics results for Investment are not as clean. To see this we can back out our investment variables σ_l and σ_h, which mechanically decrease with each of the quality variables,

$$\begin{aligned} \sigma_h &= \frac{\phi}{f}\left(\frac{f-\psi}{\phi-\psi}\right) \\ \sigma_l &= \frac{1-\phi}{1-f}\left(\frac{f-\psi}{\phi-\psi}\right) \end{aligned} \tag{6}$$

Total investment in education is therefore

$$I \equiv f\sigma_h + (1-f)\sigma_l = \frac{f-\psi}{\phi-\psi} \tag{7}$$

Which has an ambiguous sign when we take derivatives with respect to cost, productivities and even the 'real' skill premium λ.

This ambiguity is interesting nevertheless. The implication is that an increase in the cost of education might actually increase the demand for education. Why would this happen? When costs increase, skill becomes less attractive for both types of workers. However, when the low ability types retract from school, the quality of skilled workers improves. Firms are willing to pay a higher wage for a worker of higher expected ability. This increase in the value of skill is due to an increase in its value as a signal on ability. This increase in skill's relative value could potentially dominate the absolute increase in the cost of skill. Figure 5 presents such a case. Wherever costs increase investment, we have an upward sloping demand curve.

The one parameter that unambiguously affects investment is f, the ability prevalence in the population, which represents the initial endowment of human capital in the population. This is an important parameter of the economy, and has a central role in an environment with asymmetric information. Any inference on behalf of the ignorant party takes this prior ability distribution as the basis for subsequent updating. To see how worker's choice depends on this initial ability distribution, we begin with the following Lemma,

Lemma 4 *Initial human capital endowment does not affect self-selection or wages.*

Proof. The problem stated in terms of the probability parameters ϕ and ψ does not involve the fraction of able to unable persons, f. ■

Workers sort themselves in to skill to make the returns (per signal) equal across skills. This implies some relationship between the quality of workers in each skill level regardless of the initial distribution of abilities. The result on wages follows since they depend on the endogenous compositions , and not the original underlying distribution.

Investment, however, is affected by f.

Proposition 9 *Investment of both type of workers increases with initial human capital, f.*

Proof. Since f does not affect equilibrium ϕ and ψ, we only need to consider the direct effect of f on investment. Differentiating the expressions for investment [6] with respect to f we have $\frac{d\sigma_h}{df} = \frac{\phi\psi}{f^2(\phi-\psi)} > 0$ and $\frac{d\sigma_l}{df} = \frac{(1-\phi)(1-\psi)}{(1-f)^2(\phi-\psi)} > 0$ ■

There are externalities to human capital. A population which is endowed with more human capital will choose to invest even further in its human capital. Underlying this result is the complementarity between workers' choices. Consider an increase in the population's ability. Since the equilibrium fractions of able to unable workers has to be the same to keep returns equal, then high ability workers must increase their investment in skill. However, this entails an increase in investment of low ability types due to the complementarities.

This result is extremely relevant when discussing the welfare of disadvantaged groups. Even a high ability individual will invest less in education if there are fewer able individuals in his group. Since any identifiable group is subject to a separate market, we can compare what will happen to such groups that differ along the ability dimension. In a more dynamic setting where investment in education today affects the ability of the next generation, we are heading toward wage dispersion and increased inequality between groups. To break away from this course of events we need to invest in disadvantaged groups early to increase their ability to be productive participant in the labor market. In addition to the straightforward value added, there will be the additional positive information externality just discussed.

We haven't referred to any identifiable feature such as gender, race or ethetnicity. However, this is undue caution. What we call ability is really the productivity of a worker employed in a labor market with some specific technology. It is hardly controversial to claim that some groups are less productive than others: new immigrants might have some cultural or language barriers, females might have skills not suited for a predominantly male industry, etc. The implication of the model for these more concrete examples would be that the increase in female education attainment may have also been exacerbated by the increased share of females who were now well prepared to take part in market production. Their increased investment pulled into school females who were initially less prepared for market work.

4.4 Welfare

Welfare, too, has an ambiguous response to changes in prices and productivity. This follows directly from the ambiguity of investment. The welfare loss (DWL) results from the inefficiency imposed by the information friction. It is equal to the weighted sum of efficiency loss from workers investing when they should not, or not investing when they should. While we have assumed it is inefficient for the low ability workers to invest in skill (Assumption $A2$), only now do we have to specify if investment for the high type is efficient or not,

$$\begin{aligned} DWL &= (1-f)\sigma_l(C - a_l(\lambda-1)) + f(1-\sigma_h)((\lambda-1)a_h - C) \quad \text{if} \quad \lambda a_h - C > a_h \\ &= (1-f)\sigma_l(C - a_l(\lambda-1)) + f\sigma_h(C - (\lambda-1)a_h) \quad \text{if} \quad \lambda a_h - C < a_h \end{aligned} \tag{8}$$

We could have an interior equilibrium where both type of workers should not invest (and would not if information was complete,) but in equilibrium they do[19].

Note also that an increase in human capital investment is not always welfare improving. We have assumed that the investment of low ability workers is inefficient, so any investment on their part reduces welfare. Even if the investment of the high ability worker is efficient we would still need to weigh the relative loss and gain.

5 Endogenous Cost of College

We now put the model in context of the college market and think about the general equilibrium consequences of changing market conditions. Thus far, the signaling equilibrium provided the demand for education, taking the cost of college as given. The possibility that human capital investment increases with the cost of investment can create non-standard results in the market. To see the full effects we need to see how the cost of college is determined in equilibrium. We therefore specify how production of education takes place, and solve for the equilibrium tuition and quantity of students. We then discuss how the market will react to an increase in skill biased technology, a change in the college market structure, etc.

5.1 College Production

In essence tuition cost is endogenized by specifying that the production of skill uses scientists who are in limited supply. In particular, we will assume that a constant fraction of expenditures is spent on these scarse resources. This natural assumption allows for a supply curve which is not perfectly elastic. College expenditure data suggests that college production is highly labor intensive, with the share of expenditures on research and instruction going around 0.4[20].

We therefore add a higher education sector in the following straightforward way. Assume that production of college graduates L_H takes as inputs a general aggregate good Y whose price is normalized to one and some scientists $\overline{S}$ who are in limited supply and earn a competitive wage w. Production of college graduates, L_H, takes the cobb-douglas form with the share of scientists being α,

$$L_H = S^\alpha Y^{1-\alpha} \tag{9}$$

[19] This is the same as in the standard signaling environment.

[20] The Integrated Postsecondary Education Data System (IPEDS) data.

Competitive firms sell college education to students at the tuition rate of C and therefore face the standard maximization problem,

$$\max_{S,Y} CS^{\alpha}Y^{1-\alpha} - wS - p_y Y \tag{10}$$

From the first order conditions we can solve for the cost of college, C

$$C = \chi\,(w)^{\alpha} \tag{11}$$

Where $\chi = \alpha^{-\alpha}\,(1-\alpha)^{-(1-\alpha)}$.

The scientists wage, w, is given by the equilibrium in the scientists market. From firm's maximization we have the demand for scientists given by

$$S^d = L_H w^{\alpha-1}(\frac{1-\alpha}{\alpha})^{\alpha-1} \tag{12}$$

This must be equal to the fixed supply, $\overline{S}$. Substituting for the wage w from (11) we have the supply of college given by

$$L_H^s = \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha}\left(\frac{C}{\chi}\right)^{\frac{1-\alpha}{\alpha}}\overline{S} \tag{13}$$

This is a standard upward sloping supply curve which increases with the price, C. It exhibits economies of scale if the share of scientists is smaller than half ($\alpha < 0.5$).

5.2 Equilibrium in the College Market

We are ready to solve for the college market equilibrium.

Definition 5 *The college market equilibrium is given by* $\{\sigma^*, w_e^*(s), \mu^*(\cdot|e,s)\} \cup \{C\}$ *which satisfy the signaling equilibrium conditions above and the additional college market clearing condition,*

$$\sum_i p(a_i)\sigma_i = L_H^s$$

Rewriting the clearing market condition using our parameters (ϕ, ψ), we have

$$\frac{f-\psi}{\phi-\psi}(C) = \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha}\left(\frac{C}{\chi}\right)^{\frac{1-\alpha}{\alpha}}\overline{S} \tag{14}$$

Where the solutions for $\langle\phi(C), \psi(C)\rangle$ are given by the signaling equilibrium. The demand for education generally has an ambiguous slope, as we saw in our discussion of investment. If it is upward sloping there is a potential for multiple equilibria, however, only one of them is stable. In the stable equilibrium, the elasticity of demand must be greater than the elasticity of supply.

With the equilibrium in place, we can now look at the comparative statics, and show how the model explains the recent trends, and what are the predictions for selection. Consider first a skill biased technical change (λ). Selection first unambiguously declines, with a likely increase in investment. The increase in college demand increases tuition costs, which through the general equilibrium effect increase selection, and counteract the initial decline. The skill premium likely increases, both from the initial skill-biased change, and the second order increase in selection. Next consider an increase in initial human capital (f). Investment increases, with no first order effect on selection. The increase in tuition fees due to increased demand increases selection and the skill premium unambiguously.

Both of these explanations fit the broad facts of the college market: increased tuition, increased enrollment and an increased premium for education. They differ along the new dimension that the model introduces: self selection. A $SBTC$ has a complex effect on selection, which can be negative if the direct and general equilibrium effects are strong enough. An increase in human capital endowment will entail an increase in selection.

The model is consistent with the college market facts. It provides a possible mechanism that takes into account that workers have heterogenous abilities, and that their choices may cause an additional selection effect. While there is reasonable consensus that the major changes in the labor market over the past few years are due to a skill biased technical change, this model offers an alternative trigger. An exogenous increase in human capital can lead to the same observed consequences. The likelihood of such an increase in human capital is left for future research.

6 Conclusion

An equilibrium model of the college market was presented in which demand for education is part of a special signaling equilibrium of the labor market and in which the production of education uses scarce scientists as factors.

This paper contributes to the signaling literature by exploring the possibility of self-selection that is due to differential returns and not differential costs. It shows that the conditions for an equilibrium with positive selection to arise is that it is socially inefficient for low ability workers to invest in education. Solving the problem in terms of the quality variables instead of the standard quantity ones turns out to be very useful, since in an environment of imperfect information quality affects prices. Stating the problem in this way allows us to derive clean comparative statics, which do not exist for the investment variables: self selection increases with the net cost of low ability types investing in education.

Finally, the model takes a step beyond the signaling framework by looking at the general equilibrium implication of the signaling equilibrium. When the production of college education depends on scientists which are in limited supply, the equilibrium wages feed back into college tuition (and back into the equilibrium selection). An initial exogenous technological advance which biases skill will result in decreased self selection. The increase in high-skill wages, while somewhat mitigated by the decline in quality, still increases the cost of supplying college education. This increase in costs works to reverse the original decline in student quality.

The two main results of the paper have important implication for inequality and social policy. The first concerns the debate around the ongoing expanded financial assistance for education at the state and federal levels. This analysis suggests that an increase in grants, increases the relative investment of low ability workers in education. While it may be inefficient in the short run, it will mitigate inequality. This is in stark contrast to Hendel and al. (2005) who use a similar framework, augmented with credit constraints to reach an opposite conclusion.

The second result compares the education decisions of identifiable groups which differ in their average human capital. The likelihood of going to college increases for all agents in the group which has an initial higher productivity potential. Individuals with initial low ability will be pulled up to earn an education because there are more able individuals in their group. This suggests that early intervention programs that increase potential market productivity have yet another benefit. These programs create a positive externality on agents in the same group that have not been treated by the policy.

The quality of skilled and unskilled labor is an evasive empirical entity. Nevertheless, this work suggest that we may want to get a better understanding of the ability composition of skill-groups: not only to correct for such compositional bias of the 'real' skill premium, but as an object in itself. It would be valuable to see how the quality of the workforce is endogenously determined by the wages rewarded and the tuition costs, and in turn how it affects those same prices.

7 Appendix

The result that no separating equilibrium exists might be viewed as a weakness because it implies a discontinuity of the solution in the neighborhood of perfect information. To see this, assume the parameters are such that workers' optimal choice under full information is separation. As information gets better the equilibrium will converge to the fully separating equilibrium, but the limit will not exist. We solve this discontinuity and restore the existence of the separating equilibrium by small behavioral perturbations. In this way beliefs will never ignore new information completely.

Lemma 5 *(Robustness of Proposition 1) Proposition 1 is not robust to small behavioral trembles: If separation is optimal when information is complete than for any small fraction* 2ϵ *of workers of each type who randomize between the two skill levels there exists* p_{h0} *and* p_{l0} *such that for any* $p_h > p_{h0}$ *and any* $p_l < p_{l0}$ *there exists a separating equilibrium.*

Proof. For any interior beliefs $\mu(a_i|e,s)\in(0,1)$ taking the limits as $p_h\longrightarrow 1$ and $p_l\longrightarrow 0$ we have $\lim w_e(h)=\lambda a_h$ and $\lim w_e(l)=a_l$ so that $\lim Ew_H(a_h)-C>\lim Ew_L(a_h)$ and $\lim Ew_H(a_l)-C<\lim Ew_L(a_l)$ and separation is optimal. To sustain this as an equilibrium we must have the Bayesian beliefs be interior. But this is always the case with a fraction 2ϵ of agents randomizing since posteriors are updates on the interior 'interim-priors' given by: $p_H^\epsilon=p(a_h|H)=\frac{(1-\epsilon)f}{(1-\epsilon)f+\epsilon(1-f)}$ and $p_L^\epsilon=p(a_h|L)=\frac{\epsilon f}{\epsilon f+(1-\epsilon)(1-f)}$. ■

Recall the definition of the likelihood variables ('interim-priors'):

$$\begin{aligned}\phi &\equiv p(a_h|H)=\frac{f\sigma_h}{f\sigma_h+(1-f)\sigma_l}\\ \psi &\equiv p(a_h|L)=\frac{f(1-\sigma_h)}{f(1-\sigma_h)+(1-f)(1-\sigma_l)}\end{aligned}$$

We can fully express our equilibrium in terms of the four posteriors,

$$\begin{aligned}\widehat{\phi}_h &\equiv \mu(a_h|H,s=h)=\frac{\phi p_h}{\phi p_h+(1-\phi)p_l}\\ \widehat{\phi}_l &\equiv \mu(a_h|H,s=l)=\frac{\phi(1-p_h)}{\phi(1-p_h)+(1-\phi)(1-p_l)}\\ \widehat{\psi}_h &\equiv \mu(a_h|L,s=h)=\frac{\psi p_h}{\psi p_h+(1-\psi)p_l}\\ \widehat{\psi}_l &\equiv \mu(a_h|L,s=l)=\frac{\psi(1-p_h)}{\psi(1-p_h)+(1-\psi)(1-p_l)}\end{aligned}$$

Proof. of Proposition 2 : Types of Equilibria.

(i) In a completely mixed solution both types must be indifferent, $\forall i, Ew_H(a_i)-C=Ew_L(a_i)$. As shown in the text (section 2) these two equations can be rewritten as 3 and 4. These are two equations in (ϕ,ψ) in the second degree. Using brute-force and explicitly solving we get that a necessary condition for existence is $\frac{\lambda-(A+1)}{\lambda-A}\frac{A}{A+1}>\frac{4ab}{(a+b)^2}$. We still need to make sure when these solutions translate to probabilities between (0,1). (finish).

(ii) We will show that $(\sigma_h,\sigma_l)^*=(1,1)$ can be a part of an equilibrium iff $\frac{p_hp_l}{p_h+p_l}+\frac{(1-p_h)(1-p_l)}{1-p_h+1-p_l}>\frac{A}{\lambda}$. Compatible beliefs with these strategies are $\widehat{\phi}_h=\frac{p_h}{p_h+p_l};\widehat{\phi}_l=\frac{(1-p_h)}{(1-p_h)+(1-p_l)}$. Assign off equilibrium path beliefs to be $\widehat{\psi}_h=0$ and $\widehat{\psi}_l=0$. If a_l prefers H, so does a_h, because he always has higher probability of good signal. Finally, a_l prefers H iff $p_l\lambda\widehat{\phi}_h+(1-p_l)\lambda\widehat{\phi}_l>A\Longleftrightarrow\frac{p_hp_l}{p_h+p_l}+\frac{(1-p_h)(1-p_l)}{1-p_h+1-p_l}>\frac{A}{\lambda}$.

(iii) We will show that $(\sigma_h,\sigma_l)^*=(0,0)$ can always a part of an equilibrium. Compatible beliefs are $\widehat{\psi}_h=\frac{p_h}{p_h+p_l};\widehat{\psi}_l=\frac{(1-p_h)}{(1-p_h)+(1-p_l)}$. Assign off path beliefs to be $\widehat{\phi}_h=0$ and $\widehat{\phi}_l=0$. If a_l prefers H, so does a_h, because he always has higher probability of good signal. a_l prefers H if $p_l(-\widehat{\psi}_h)+(1-p_l)(-\widehat{\psi}_l)<A$, $\Leftrightarrow\frac{p_lp_h}{p_h+p_l}+\frac{(1-p_l)(1-p_h)}{(1-p_h)+(1-p_l)}>-A$ (which is always true if $A>0$).

(iv) We will show that $\sigma_h^*=1;\sigma_l^*\in(0,1)$ can be a part of an equilibrium if there is a solution $\sigma_l\in(0,1)$ which solves $p_l\frac{p_h}{p_h+\sigma_lp_l}+(1-p_l)\frac{(1-p_h)}{(1-p_h)+\sigma_l(1-p_l)}=\frac{A}{\lambda}$. The compatible beliefs are given by $\widehat{\phi}_h = \frac{p_h}{p_h+\sigma_l p_l}; \widehat{\phi}_l = \frac{(1-p_h)}{(1-p_h)+\sigma_l(1-p_l)}$; $\widehat{\psi}_h = 0$; $\widehat{\psi}_l = 0$ If a_l is indifferent, a_h will surely prefer H. We therefore only require $a_l's$ indifference condition to hold, $p_l\left(\lambda\widehat{\phi}_h\right) + (1-p_l)\left(\lambda\widehat{\phi}_l\right) = A$ which is the condition for σ_l given.

(v) We will show that $\sigma_l^* \in (0,1)$, $\sigma_h^* = 0$ can be a part of an equilibrium only if $A < 0$. The compatible beliefs are $\widehat{\phi}_h = 0$, $\widehat{\phi}_l = 0$, $\widehat{\psi}_h = \frac{p_h}{p_h+(1-\sigma_l)p_l}$, $\widehat{\psi}_l = \frac{(1-p_h)}{(1-p_h)+(1-\sigma_l)(1-p_l)}$. To have a_l indifferent requires $p_l\left(-\widehat{\psi}_h\right) + (1-p_l)\left(-\widehat{\psi}_l\right) = A$ which can only be true for $A < 0$. ■

Proof. of Proposition 3 (Pooling is not Divine): We will show that pooling on L fails the divine criterion if both workers prefer skill over the equilibrium pooling, when firms believe only able persons acquire skill. For $i = h, l$ define $g^i : [0,1] \longrightarrow [0,1]$, $g^i(x) \equiv p_i \frac{xp_h}{xp_h+(1-x)p_l} + (1-p_i)\frac{x(1-p_h)}{x(1-p_h)+(1-x)(1-p_l)}$. Note that $g^i(x)$ is increasing and monotone in x, with $g^i(0) = 0$ and $g^i(1) = 1$, and that $g^h(x) > g^l(x)$ because of monotone likelihood assumption. The conditions of the proposition state that $\lambda g^i(1) > g^i(f) + A$. Pooling on L implies $\lambda g^i(0) < g^i(f) + A$. By the intermediate value theorem there exist x^h and x^l s.t. $\lambda g^i(x^i) = g^i(f) + A$. Because $g^h(f) > g^l(f)$ we have $g^h(x^h) > g^l(x^l)$. Except for non generic payoffs this implies $x^h \neq x^l$. Assume $x^i > x^j$ then the for all $x_0 \in (x^j, x^i)$ we have $\lambda g^i(x_0) < g^i(f) + A$ but $\lambda g^j(x_0) > g^j(f) + A$. Hence $D(a_i, H) \cup D^0(a_i, H) = [x^i, 1]$ and $D(a_j, H) = (x^j, 1]$ with $D(a_i, H) \cup D^0(a_i, H) \subset D(a_j, H)$.

Similarly, pooling on H fails the divine criterion if both workers prefer no-skill over pooling on H, when firms believe only able persons don't acquire skill. ■

For the proof of proposition 6 we will use the shortcut notation $\widetilde{p}$ for the probability of having a high signal conditional on being in the H group, and $\widetilde{q}$ for the probability of having a high signal conditional on being in the L group, and finally $\widetilde{f}$ as the unconditional probability, That is

$$\begin{aligned}
\widetilde{p} &= p(s=h|H) = \phi p_h + (1-\phi)p_l \\
\widetilde{q} &= p(s=h|L) = \psi p_h + (1-\psi)p_l \\
\widetilde{f} &= p(s=h) = fp_h + (1-f)p_l
\end{aligned}$$

Lemma 6 $A2 \Longrightarrow \phi > f > \psi \Longleftrightarrow \widetilde{p} > \widetilde{f} > \widetilde{q}$

Proof. of Lemma: By Proportions 2: $A2 \Longrightarrow \phi > \psi \Longleftrightarrow \sigma_h > \sigma_l$. Hence $\phi = \frac{f\sigma_h}{f\sigma_h+(1-f)\sigma_l} > f$ and $\psi = \frac{f(1-\sigma_h)}{f(1-\sigma_h)+(1-f)(1-\sigma_l)} < f$. From $\psi < f < \phi \Longleftrightarrow \psi p_h + (1-\psi)p_l < fp_h + (1-f)p_l < \phi p_h + (1-\phi)p_l$ since $p_h > p_l$ ■

Proof. of Proposition 6 (self-selection): By implicitly differentiating the equilibrium equations 3 and 4 we get $\frac{d\psi}{dC} = -\frac{1}{\lambda(a_h-a_l)p_h p_l(1-p_h)(1-p_l)} \frac{\left(\frac{\lambda p_h p_l}{\widetilde{p}^2}\right)-\left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{p})^2}\right)}{\left(\frac{1}{\widetilde{p}^2}\right)\left(\frac{1}{(1-\widetilde{q})^2}\right)-\left(\frac{1}{(1-\widetilde{p})^2}\right)\left(\frac{1}{\widetilde{q}^2}\right)}$ The denominator is negative by the Lemma. The Nominator is negative since $\frac{p_h p_l}{(1-p_h)(1-p_l)} < \frac{\widetilde{f}^2}{(1-\widetilde{f})^2}$ by assumption and $\frac{\widetilde{f}^2}{(1-\widetilde{f})^2} < \frac{\widetilde{p}^2}{(1-\widetilde{p})^2}$ by the Lemma. Similarly differentiating for ϕ and subtracting we have $\frac{d(\phi-\psi)}{dC} > 0 \iff \left(\frac{(1-p_h)(1-p_l)}{(1-\widetilde{q})^2}\right) - \left(\frac{p_h p_l}{\widetilde{q}^2}\right) + \left(\frac{\lambda p_h p_l}{\widetilde{p}^2}\right) - \left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{p})^2}\right) < 0$. But by the Lemma this expression is smaller than $\left(\frac{(1-p_h)(1-p_l)}{(1-\widetilde{f})^2}\right) - \left(\frac{p_h p_l}{\widetilde{f}^2}\right) + \left(\frac{\lambda p_h p_l}{\widetilde{f}^2}\right) - \left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{f})^2}\right) = \left(\frac{p_h p_l}{\widetilde{f}^2} - \frac{(1-p_h)(1-p_l)}{(1-\widetilde{f})^2}\right)(\lambda - 1) < 0$ by assumption. All of the other parameters, except λ, have exactly the same expression with only the leading term changing a bit with the appropriate signs. So we only need to check the derivative with respect to λ. This is messy, but turns out the same condition is sufficient for $\frac{d(\phi-\psi)}{d\lambda} < 0$: ∎

Proof. of Proposition 7 (skill premium): Adding the appropriate λ to the proof of proposition(6) we have $\frac{d(\lambda\phi-\psi)}{dC} > 0 \iff \left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{q})^2}\right) - \left(\frac{\lambda p_h p_l}{\widetilde{q}^2}\right) + \left(\frac{\lambda p_h p_l}{\widetilde{p}^2}\right) - \left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{p})^2}\right) < 0$ which is true by the Lemma proceeding Proposition(6). There is no need for condition1. ∎

8 Bibliography

References

[1] Acemoglu, Daron, "A Microfoundation for Social Increasing Returns in Human Capital Accumulation," *Quarterly Journal of Economics*, August 1996, v. 111(3), pp. 779-804

[2] ———, "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence," *American Economic Review*, December 1999, v. 89(5), pp. 1259-78

[3] ———, "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, March 2002, v. 40(1), pp. 7-72.

[4] Angrist, Joshua D.; Krueger, Alan B. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, November 1991, v. 106, iss. 4, pp. 979-1014

[5] Arrow, Kenneth J., "Higher Education as a Filter," *Journal of Public Economics*, July 1973a, v. 2(3), pp. 193-216

[6] ———, "The Theory of Discrimination," in Orley Ashenfelter and Albert Reis eds., *Discrimination in Labor Markets*, Princeton, NJ.: Princeton University Press, 1973b, pp. 3-33.

[7] Ashenfelter, Orley; Rouse, Cecilia, "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins," *Quarterly Journal of Economic*s, February 1998, v. 113, iss. 1, pp. 253-84

[8] Autor, David H and Katz, Lawrence F., "Changes in the Wage Structure and Earnings Inequality." *Handbook of labor economics.* 1999 (3A) 1463-1555. Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

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

[10] Autor, David H., Katz, Lawrence F. and Kearney, Melissa S., "Rising Wage Inequality: The Role of Composition and Prices". 2005, *National Bureau of Economic Research*, NBER Working Papers: 11628.

[11] Autor, David H., Levy, Frank and Murnane, Richard J., "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics*, November 2003, v. 118(4), pp. 1279-1333

[12] Banks, Jeffrey S. and Sobel, Joel, "Equilibrium Selection in Signaling Games," *Econometrica*, May 1987, v. 55, iss. 3, pp. 647-61

[13] Bedard, Kelly, "Human Capital versus Signaling Models: University Access and High School Dropouts," *Journal of Political Economy*, August 2001, v. 109, iss. 4, pp. 749-75

[14] Berman, Eli; Bound, John; Griliches, Zvi, "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures," *Quarterly Journal of Economics*, May 1994, v. 109(2), pp. 367-97

[15] Blackburn, McKinley L.; Neumark, David, "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look," *Review of Economics and Statistics*, May 1995, v. 77, iss. 2, pp. 217-30

[16] Cameron, Stephen V. and Heckman, James J., "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males" *Journal of Political Economy*, April 1998, v. 106(2), pp. 262-333.

[17] Cameron, Stephen V. and Taber, Christopher, "Estimation of Educational Borrowing Constraints Using Returns to Schooling" *Journal of Political Economy*, Part 1 Feb. 2004, v. 112(1), pp. 132-82.

[18] Card, David, "The Causal Effect of Education on Earnings" in *Handbook of labor economics*. 1999(3A) 1801-63, . Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

[19] Card, David; Lemieux, Thomas, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, May 2001, v. 116, iss. 2, pp. 705-46

[20] Cho, In-Koo; Kreps, David M., "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, May 1987, v. 102(2), pp. 179-221

[21] Coate, Stephen and Loury, Glenn C., "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, December 1993, v. 83(5), pp. 1220-40.

[22] Feldstein, Martin., "College Scholarship Rules and Private Saving." *American Economic Review*, June 1995, v. 85(3), pp. 552-66.

[23] Goldin, Claudia and Katz, Lawrence F., " The Returns to Skill in the United States across the Twentieth Century." 1999, *National Bureau of Economic Research*, NBER Working Papers: 7126.

[24] Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, Jan. 1977, v. 45(1), pp. 1-22

[25] Hendel, Igal, Shapiro, Joel and Willen, Paul., "Educational Opportunity and Income Inequality" *Journal of Public Economics*, June 2005, v. 89(5-6), pp. 841-70.

[26] Hoxby, Caroline M., "How the Changing Market Structure of U.S. Higher Education Explains College Tuition." 1997, *National Bureau of Economic Research*, NBER Working Papers: 6323.

[27] Hvide, Hans K., "Education and the Allocation of Talent," *Journal of Labor Economics*, Special Issue October 2003, v. 21, iss. 4, pp. 945-76

[28] Juhn, Chinhui, Kim, Dae Il and Vella, Francis., "The Expansion of College Education in the United States: Is There Evidence of Declining Cohort Quality?" *Economic Inquiry*, April 2005, v. 43(2), pp. 303-15.

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

[30] Krueger, Alan B., "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989," *Quarterly Journal of Economics*, February 1993, v. 108, iss. 1, pp. 33-60

[31] Krusell, Per, Ohanian, Lee E., Rios-Full, Jose-Victor and Violante Giovanni L., "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," Econometrica, September 2000, v. 68, iss. 5, pp. 1029-53

[32] Lang, Kevin; Kropp, David, "Human Capital versus Sorting: The Effects of Compulsory Attendance Laws," *Quarterly Journal of Economics*, August 1986, v. 101(3), pp. 609-24

[33] Moro, Andrea and Norman, Peter., "A General Equilibrium Model of Statistical Discrimination," *Journal of Economic Theory*, January 2004, v. 114(1), pp. 1-30

[34] Murnane, Richard J.; Willett, John B.; Levy, Frank, "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics*, May 1995, v. 77, iss. 2, pp. 251-66

[35] Phelps, Edmund S., "The Statistical Theory of Racism and Sexism," *American Economic Review*, Sept. 1972, v. 62(4), pp. 659-61

[36] Riley, John G., "Informational Equilibrium," *Econometrica*, March 1979, v. 47(2), pp. 331-59

[37] Rosen, Sherwin and Willis, Robert J., "Education and Self-Selection." *Journal of Political Economy*, Part 2, Oct. 1979, v. 87(5), pp. S7-36

[38] Rothschild, Michael; White, Lawrence J., "The Analytics of the Pricing of Higher Education and Other Services in Which the Customers Are Inputs," *Journal of Political Economy*, June 1995, v. 103(3), pp. 573-86

[39] Roy, A. D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, III (1951), pp. 235-246

[40] Spence, A. Michael., "Job Market Signaling." *Quarterly Journal of Economics*, Aug. 1973, v. 87(3), pp. 355-74.

[41] Steinberger, Michael,. "Educational Sorting and the Returns to College: 1977-1999." Unpublished manuscript.

[42] Taubman, Paul J. and Wales, Terence J. "Higher Education, Mental Ability, and Screening." *Journal of Political Economy*, Jan.-Feb. 1973, v.

81(1), pp. 28-55..

[43] Tyler, John H., Murnane, Richard J., Willett, John B.,"Estimating the Labor Market Signaling Value of the GED" *Quarterly Journal of Economics*, May 2000, v. 115(2), pp. 431-68

[44] Weibull, Jorgen W., *Evolutionary game theory*, 1995, pp. xv, 265, Cambridge and London: MIT Press

[45] Weiss, Andrew, "A Sorting-cum-Learning Model of Education," *Journal of Political Economy*, June 1983, v. 91(3), pp. 420-42

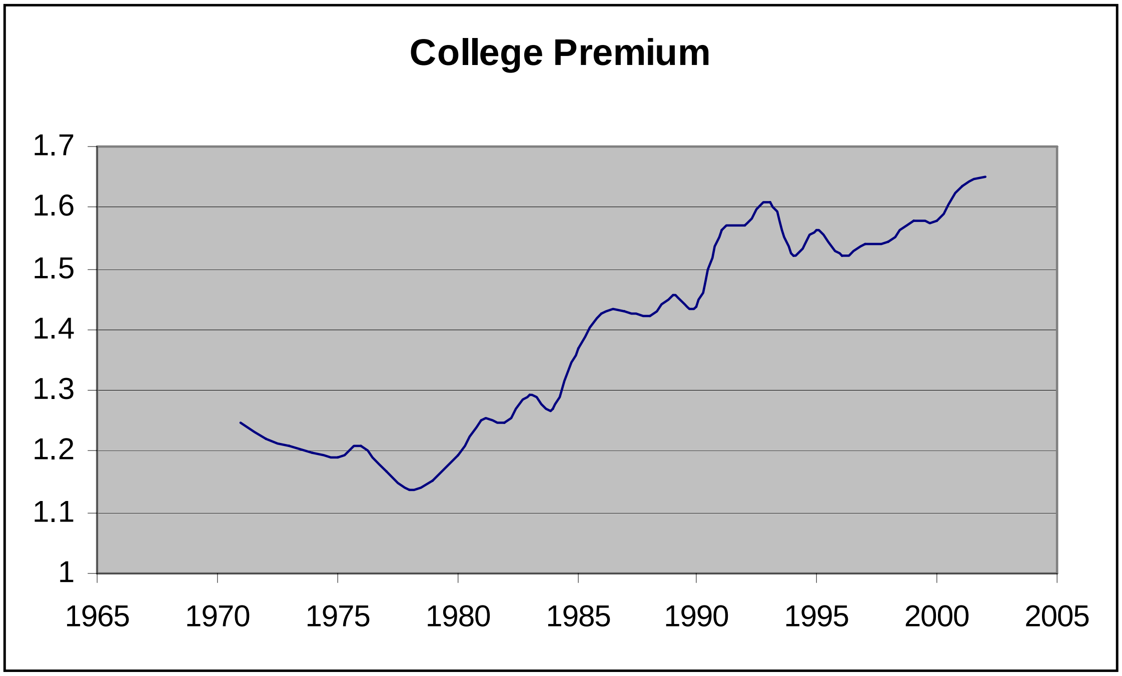


Figure 1:

Source: National Center for Education Statistics. (14-1). Median annual earnings (in constant 2002 dollars) of all full-time, full-year wage and salary workers ages 25–34, by sex and educational level: 1971–2002.

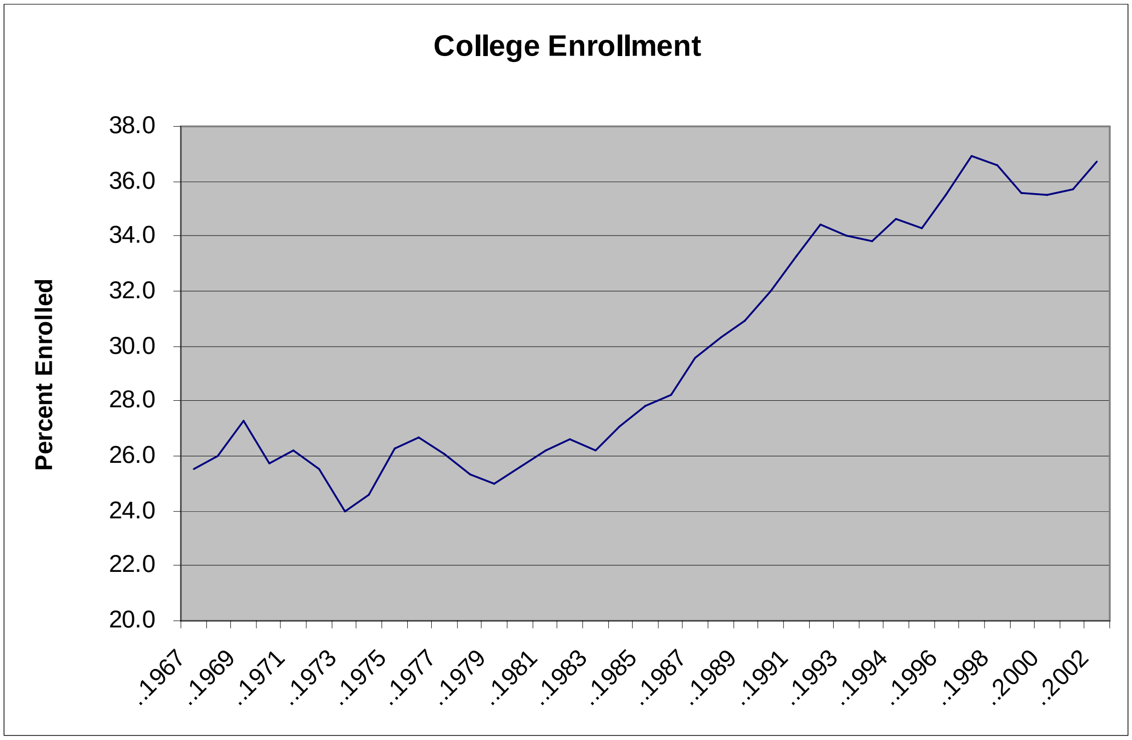


Figure 2:

Source: CPS, Percent of the population 14 to 24 years old enrolled in college

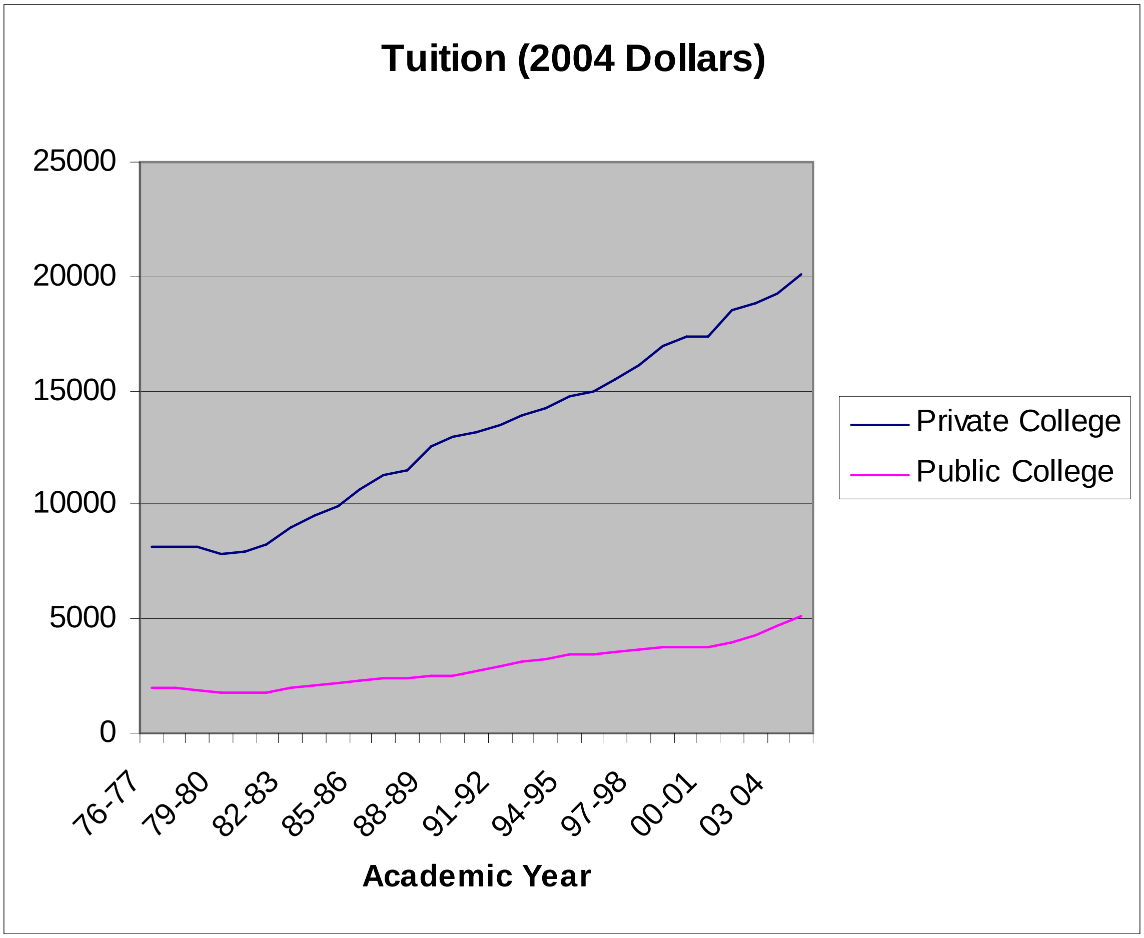
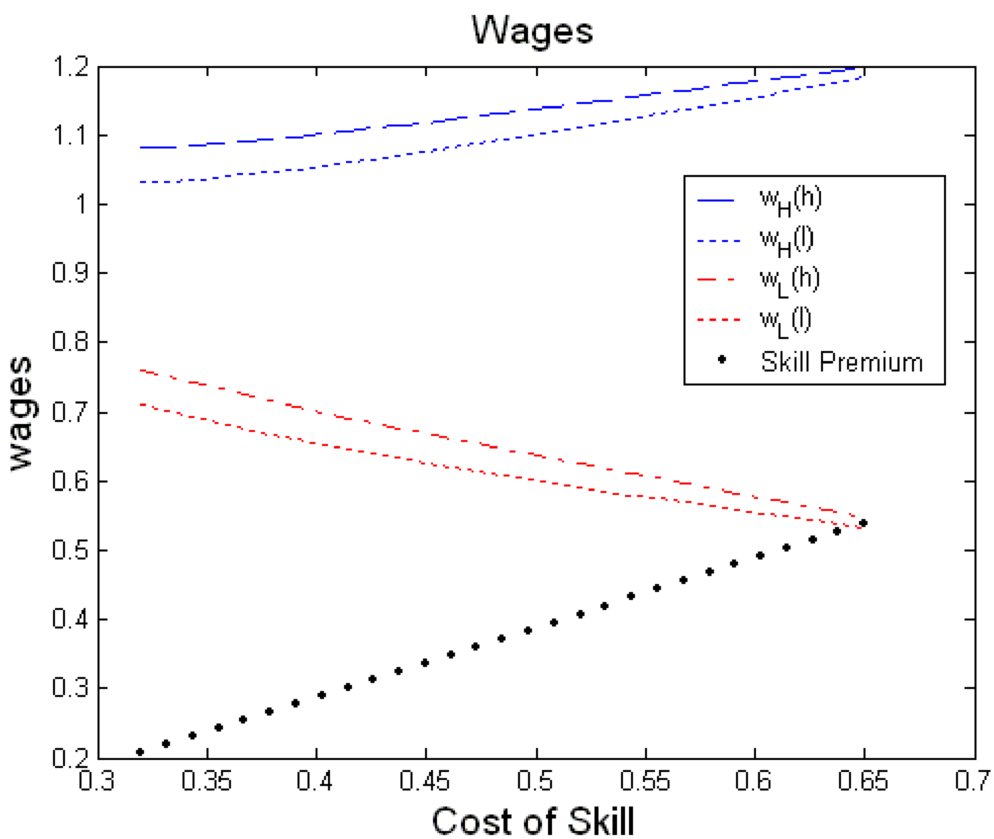


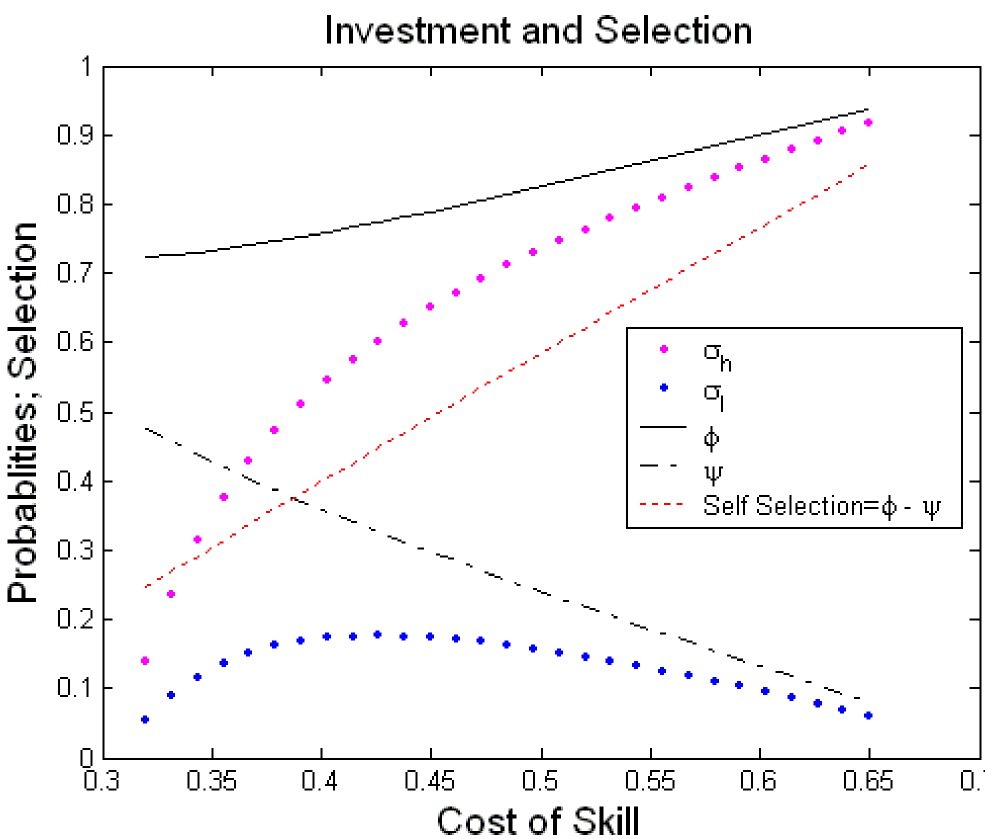
Figure 3:

Source: 1987-88 to 2004-05: data from Annual Survey of Colleges, The College Board, New York, NY, weighted by full-time undergraduate enrollment; 1976-77 to 1986-87: data from Integrated Postsecondary Education Data System (IPEDS), U.S. Department of Education, National Center for Education Statistics, weighted by full time equivalent undergraduate enrollment



Wages
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
wages
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
Cost of Skill
$w_H(h)$
$w_H(l)$
$w_L(h)$
$w_L(l)$
Skill Premium

Figure 4:



Investment and Selection
Probablities; Selection
Cost of Skill
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
σ_h
σ_l
ϕ
ψ
Self Selection=ϕ - ψ

Figure 5: