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

Tali Regev*

October 26, 2009

Abstract

This paper introduces and explores signalling in the market for education based on heterogeneity in the returns to education rather than heterogeneity in costs. Workers of heterogeneous abilities face the same costs, yet a larger proportion of able individuals self-select to attend college since they are more likely to get higher returns. With imperfect information, the skill premium is an outcome which depends on the equilibrium quality of college attendees and non attendees. Incorporating a production function of college education, I discuss the properties of the college market equilibrium. A skill-biased technical change directly decreases self-selection into college, but the general equilibrium effect may overturn the direct decline, since increased enrollment and rising tuition costs increase self-selection. Higher initial human capital has an external effect on subsequent investment in school: All agents increase their education, and the higher equilibrium tuition costs increase self-selection and the college premium. This model can help explain the steady trends in increasing tuition costs, college enrollment, and the college wage gap through its relationship to the quality of college graduates. It suggests that the signaling role of education might be an important yet largely neglected ingredient in these recent changes. JEL codes: J24, I21, J31

*Email talir@tau.ac.il. I am grateful to my advisors Daron Acemoglu and Ivan Werning for their helpful suggestions and support. I also thank David Autor, Karna Basu, Ricardo Caballero, Mary Daly, Mike Golosov, Guido Lorenzoni, Hendricks Lutz, Veronica Rappoport, Eric Swanson, Anita Todd, Rob Valletta, Yoram Weiss, Daniel Wilson and seminar participants at Bar Ilan, Ben Gurion, Boston, the Federal Reserve Bank of San Francisco, Haifa, Hebrew University, Massachussetts Institute of Technology and Tel-Aviv for their comments. Jeremy Gerst provided excellent research assistance.

1 Introduction

The wage gap between college graduates and high school graduates has dramatically increased since the late 1970s as a growing number of college graduates has entered the job market (see Figure 1).[1] During the same period, the cost of college has jumped more than twofold. This paper explores how these trends are related to the role of college education as a signal of worker ability to potential employers. Following recent evidence which suggests this signaling component to education could be sizeable yet untapped, I construct a new signaling model of the demand for skills, skill premiums, and college market equilibrium.

This paper departs from the standard signaling framework (Spence 1973) by focusing on returns to education rather than costs. I assume workers face the same costs, but can expect different returns from education since employers are likely to perceive their worker's ability.[2] Allowing an employer to have some knowledge of a worker's ability introduces a form of single-crossing property, which is responsible for the worker's self-selection. Nevertheless, since the information available to firms is limited, they still need to infer a worker's initial ability by using the composition of his education group, thus preserving the value of education as a signal. One reason I focus on these differing returns rather than costs is the problematic assumption in the literature that workers' abilities are negatively correlated with their costs of attending college. It seems unlikely that low-ability workers face substantially higher costs than more able workers. On the contrary, if a large part of the cost of going to college is forgone earnings, then correlation between ability and costs could even be positive.[3] Moreover, setting up a signaling model in which self-selection arises because of employer learning about productivity, refutes Lang and Topel's (2004) critique of cost-based signaling model, as they argue that the strong evidence for employer learning limits the value of schooling as a signal of unobserved ability.

As another contribution to the literature, this special signaling model allows a totally mixed equilibrium for college choice. It is reasonable to focus on this mixed equilibrium since a separating equilibrium does not exist, and hence the Riley (1979) critique does not apply.[4] This mixed equilibrium is not only plausible from an empirical viewpoint, but it has the additional benefit that self-selection to education—and hence the skill premium—does not depend solely on the exogenous distribution of initial abilities in the population.

To close the college market, I introduce a college supply function into the model. I assume

[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 (2007), Autor and Katz (1999), Autor, Katz and Kearney (2005), and Card (1999).

[2] See also Arrow (1973b) for educational signaling. See Weiss (1983) and Hvide (2003) for a setup similar to mine, but where workers don't know their abilities.

[3] A point stressed by Griliches (1977) for instance. Altonji (1993) provides evidence that the returns from post-secondary education are higher for more able workers.

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

college production uses some scientists who are in limited supply and whose wage is determined in equilibrium. Combining this with the demand for college, I solve for the equilibrium in the college market. I use this framework to describe the recent changes in the college market with the augmented selection prediction.

I show that an equilibrium with positive ability selection into skill arises when it is socially inefficient for the low-ability worker to invest in schooling. I proceed to show that self-selection increases when the cost of college is higher or when the productivity of college graduates is lower. Finally I show that total investment in education may actually increase with college costs. Behind this result are strategic externalities. Lower net returns to college make schooling less attractive for all workers. When low-ability workers shift away from college, the signaling value of going to 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 slope upwards.

The model also predicts human capital externalities. All workers invest more in their own 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. A distinct population with a larger proportion of high ability individuals will invest more in education as a society; this implies a dynamic divergence, which can be applied to differences between distinct groups based on such observables as gender, nationality or race. This result is reminiscent of the statistical discrimination literature.[5] 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.[6]

Moving on to the full college market equilibrium, I discuss how some exogenous factors affect both enrollment and the returns to college, taking into account the endogenous self-selection equilibrium. An increase in initial human capital has no first-order effect on self-selection; however, as more individuals demand education, the price of college increases, resulting in more self-selection, lower wages for less-skilled workers, and a higher college premium. More complex are the effects when a skill-biased technical change (SBTC)—which is widely believed to have taken place during the period I study—shifts demand toward more highly skilled workers relative to less-skilled workers.[7] The direct increase in the skill premium following an SBTC is undermined by the lower quality of workers who now choose to become skilled. Within-skill

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

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

[7] 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 et 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 1970's.

(residual) inequality increases, while the college premium declines, as was the case throughout the 1970s. When college enrollment and tuition increase, the general equilibrium can overturn the initial decline in self-selection, resulting in an increase in the skill premium. Such higher levels of self-selection could account not only for the wage increase for high-skill workers but also for the reduced wages of low-skill workers, a fact which otherwise remains a puzzle.[8]

There is evidence supporting the education signaling hypothesis. Many studies find a strong diploma effect, which indicates there is value in education as a signal of ability. For instance, Tyler, Murnane, and Willett (2000) find that a General Educational Development diploma signal increases wages by 10 to19 percent net of human capital effects.[9] Lang and Kropp (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. These two studies clearly show that schooling decisions are not only dependent on own schooling costs but rather on the equilibrium distribution of abilities across schooling levels.

Almost all studies find a positive selection bias.[10] 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 that new cohorts have higher returns but do not interpret this as an increase in the ability component. Murnane, Willett, and Levy (1995) find an increase in the ability composition of educated workers.[11] However, these studies are primarily concerned with the returns to skill and hence estimate only the composition of the skilled labor force, whereas this 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

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

[9] See also Jaeger and Page (1996). However, these diploma effects could also be present because individuals learn about their productivity while in school and can opt to drop out.

[10] 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 ordinary least square estimate with the unbiased instrumental variables estimate of the returns to skill, and the size of the estimated bias depends on institutional change used as the instrument. A different identification is given in Ashenfelter and Rouse (1998) who use a sample of identical twins to estimate a small upward ability bias. See Card (1999) for a complete survey.

[11] 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 50 years (for the cohorts born in 1916 to those born in 1963). Card and Lemieux (2001) interpret their findings as arising from complementarities between cohorts. Murnane, Willett, and Levy (1995) use direct test scores measures to control for the ability bias.

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, in the effect of tuition subsidies (Feldstein, 1995). Hendel, Shapiro, and Willen (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 2 presents the model and solves for the signaling equilibria. Section 3 analyses the mixed strategy equilibria. Section 4 discusses self-selection, the skill premium, the investment decision of workers and welfare. Section 5 introduces the college production and solves for the equilibrium in the college market. Section 6 concludes.

2 Model

To set up the model, first look at what enters into an individual's (a.k.a worker's) decision to acquire a costly education. Workers have a range of (initial) abilities, which they take into account when deciding whether to acquire more skills through college education. A worker's wages are determined in a competitive labor market where firms observe the worker's schooling choice and an additional proxy on a worker's initial ability. I assume for simplicity that workers of various abilities and skill levels are perfect substitutes. This simplifying assumption makes equilibrium wages depend, in effect, only on the quality of workers with the same observed skill level and leaves out the standard quantity effects.[12]

2.1 Setup

The economy consists of a continuum of mass M of firms, and a unit mass of risk-neutral workers with heterogeneous abilities. I 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 $w : E \times S \longrightarrow R^+$.

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

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.[13]

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

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 1 *(Monotone Likelihood):* $\frac{p_h(s)}{p_l(s)}$ *increases with* s*.*

The assumption that signals depend only on the worker's ability and not on his skill choice is made for simplicity.[14] Throughout, lower case $\{h, l\}$ denotes abilities, while education levels are denoted by upper case $\{H, L\}$. Since $\sigma_i(H) + \sigma_i(L) = 1$, I will use $\sigma_i \equiv \sigma_i(H)$ to denote the probability that an ability type i goes to college, wherever possible.

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:*

   $w_e^*(s) = \lambda_I \sum_i \mu^*(a_i|e, s)a_i.$

3. *Posterior beliefs are Bayesian wherever possible:*

   $\mu^*(a_i|e, s) = p(a_i)\sigma_i^*(e)p(s|a_i)/\left(\sum_{i'} p(a_{i'})\sigma_{i'}^*(e)p(s|a_{i'})\right)$

   if $\sum_{i'} p(a_{i'})\sigma_{i'}^*(e)p(s|a_{i'}) > 0$, *and any probability distribution over* $\{a_h, a_l\}$ *otherwise.*

[13] Skill could affect workers' productivities differentially, i.e., have $\lambda_l \lesseqgtr \lambda_h$ without any substantial changes to the results.

[14] 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.

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 that allows workers' 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 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 I am interested in a non-trivial composition of skills, I am 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 shown below.

To briefly characterize the full equilibrium possibilities, 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 likelihood assumption (1), beliefs are an increasing function of the signal, s, that is, for all equilibrium beliefs formed by the 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. ■

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 $\widetilde{e}$, $Ew_{\widetilde{e}}(a_l) \leqslant Ew_{\widetilde{e}}(a_h) \leqslant Ew_e(a_h) = Ew_e(a_l) = w_e$ 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 were an equilibrium in which the high-type reveals himself in e, then the Bayes rule would dictate that firms believe a worker is of 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.[15]

For the sake of completeness I 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* $\frac{p_h p_l}{p_h+p_l} + \frac{(1-p_h)(1-p_l)}{(1-p_h)+(1-p_l)} > \frac{A}{\lambda}$.

(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 $\sigma_l \in (0,1), \sigma_h = 1$ *exists 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 $\sigma_l \in (0,1), \sigma_h = 0$ *exists iff there is 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, interpreted 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 is a cost associated with low-ability types 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.

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

To be concrete, the pooling equilibria fail the divinity criterion (Banks and Sobel, 1987).[16] According to the divinity criterion, an equilibrium can be deleted if there are beliefs regarding off-the-equilibrium-path skill choice for which only one type of worker would like to deviate. To adapt the divinity criterion to this setting with minimal notations, 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 $D^0(a_i, e)$ as 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 such that $D(a_i, e) \cup D^0(a_i, e) \subset D(a_j, e)$.*

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. ■

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's rule. Consider the semi-separating equilibrium. If $A > 0$ there is 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.

I 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, 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 I find at most two mixed-strategy equilibria, which are discussed below.

[16] 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 the case discussed here, 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.

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 proportions of able workers in each skill group can be interpreted as the endogenous quality of the two groups. When firms make their wage decision, they take into account these variables, indicative of the group's composition. Then they update this "interim-prior" based on the additional individual ability signal. 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, I 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, I 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,

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

Plug the expressions for wages and 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. Use the change of variables to simplify the equilibrium equations. After some algebra, the equilibrium conditions are,

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

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

Each equation corresponds to an ability signal: equation (3) equates the rewards for a high signal ($s = h$) from getting skill or remaining unskilled, and equation (4) 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 skill levels. 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 impact 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 introduces a way to differentiate the returns of the two types, and substantially reduces the number of equilibria.

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 I explore the implication of the equilibrium allocation and wages for the objects of interest. I show that self-selection arises 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. I show how self-selection translates directly to the skill premium, look at investment and how it responds to price changes in the economy, show how investment increases in the initial level of human capital, and finally end 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 workers are more concentrated in the H sector. I 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 $\phi - \psi$.*

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 is hereafter assumed,

Assumption 2 *: 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, assume this cost is positive, that is, if information were 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 I now prove:

Proposition 5 *Self-selection arises iff it is inefficient for the lowest type to invest in skill, 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) I 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 2 basically assures 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 assumption) is sufficient for self-selection since it assures that the high-ability worker gets higher returns from skill. However, recall that Assumption 1 implies that the high-ability worker gets higher returns on any skill level (see Lemma 3).

I could also assume that it is efficient for the high-ability type to invest in skill ($a_h < \lambda a_h - C$). However, I 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 the next result on the comparative statics of self-selection, the following condition on the parameters is needed:

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

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 1 holds.[17]

Proof. See Appendix. ■

Anything that increases the costs of skill for the low-ability worker reduces his relative investment and increases self-selection. The empirical implications are straightforward. If indeed 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 ongoing increase in education costs increases self-selection, that is, the demand for college by quality applicants increases with tuition costs.

4.2 Skill Premium

I will now define and characterize the behavior of the skill premium in the model. First, note that wages in this economy depend on the composition of the skill group through an information externality. Equilibrium wages are constructed as the expected productivity of an individual with a certain skill and ability signal. Workers benefit from an increase in the quality of their skill group regardless of the specific signal they eventually emit:

Lemma 3 *A worker's wage increases with the quality of his skill-group.*

Proof. The claim is that $w_e^*(s)$ increases with $p(a_h|e)$ for all s.This is true by construction of $w_e^*(s)$. ■

As in the standard signalling model, both low and high ability workers benefit from the quality of their skill-group. Average wages rise not simply because of averaging over a group with a higher share of high paid high ability workers. Rather, the higher average wage comes from increased wages for all workers in the skill group. The composition of the group actually affects all prices, and is not a phantom "composition effect" which can be control for in wage regressions.

With this in mind, it is natural to define the skill premium as the difference between expected (or average) wages of a skilled and unskilled workers, 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 $Ew_H - Ew_L$.*

[17] This condition is only a convenient sufficient condition to prove this result. Rigorous simulations suggest the result holds under much narrower restrictions.

The skill premium increases with selection, provided that the quality of low skilled worker did not decline too much:

Proposition 7 *The skill premium increases with selection, provided that* $d\psi > -\frac{\lambda}{\lambda-1} d(\phi - \psi)$*.*

Proof. Writing out the expressions for expected wages results in $Ew_H = \lambda\left(\phi a_h + (1-\phi)a_l\right)$ and $Ew_L = \psi a_h + (1-\psi)a_l$ so that the skill premium, $Ew_H - Ew_L = (\lambda\phi - \psi)(a_h - a_l) + a_l(\lambda - 1)$, increases with $\phi - \psi$ if and only if $d\psi > -\frac{\lambda}{\lambda-1} d(\phi - \psi)$. ■

Some conditions were needed in order to resolve ambiguity in the effect of parameters on selection (Proposition 6) or the effect of selection on the wage premium (Proposition 7). However, when we look directly into the effect of parameters on the skill premium, we can prove the two following unconditional results:

Proposition 8 (*i*) *Higher investment costs,* $C \uparrow$*, increase the skill premium .*

(*ii*) *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 the change in qualities* ϕ *and* ψ*.*

Proof. See Appendix. ■

The first 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 abilities or wages, turn out to have an effect on wages. This result highlights an empirical implication of the model. Selection cannot be controlled for when looking for the skill premium because selection is part of what drives wages. The second result suggests that a skill-biased technical change would have created a larger wage dispersion without the endogenous quality adjustments. 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.

4.3 Investment in Human Capital

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

$$\begin{aligned} \sigma_h &= \frac{\phi}{f}\left(\frac{f-\psi}{\phi-\psi}\right) \\ \sigma_l &= \frac{1-\phi}{1-f}\left(\frac{f-\psi}{\phi-\psi}\right). \end{aligned} \tag{5}$$

Total investment in education is therefore

$$I \equiv f\sigma_h + (1-f)\sigma_l = \frac{f-\psi}{\phi-\psi}, \tag{6}$$

which has an ambiguous sign when I 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 the relative value of skill could potentially dominate the absolute increase in the cost of skill. Figure 3 presents such a case. Wherever costs increase investment, the demand curve slopes upwards.

The one parameter that unambiguously affects investment is f, the ability prevalence in the population, which represents the initial endowment of human capital. 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 a worker's choice depends on this initial ability distribution, 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 into skill levels 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 effect on wages follows, since wages depend on the endogenous compositions and not on the original underlying distribution.

Investment, however, is affected by f, the fraction of high ability individuals in the population.

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

Proof. Since f does not affect equilibrium ϕ and ψ, I only need to consider the direct effect of f on investment. Differentiating the expressions for investment (5) with respect to f results in $\frac{d\sigma_h}{df} = \frac{\phi\psi}{f^2(\phi-\psi)} > 0$ 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 endowed with more human capital will choose to invest even further in its human capital. Underlying this result are complementarities between workers' choices. Consider an increase in the population's ability. Since the equilibrium fraction 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, I 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, the market is heading toward wage dispersion and increased inequality between groups. One way to break away from this course of events is to invest in disadvantaged groups early to increase their ability to be productive participants in the labor market. In addition to the direct value added, there will be the additional positive information externality just discussed.

I haven't referred to any feature such as gender, race, or ethnicity as an identifying feature of a group. However, this is undue caution; what I 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 not have suitable skills for a predominantly male industry, etc. The implication of the model for these more concrete examples would be that the increase in female educational 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 dead wight 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 I have assumed it is inefficient for the low-ability workers to invest in skill (Assumption 2), only now do I have to specify whether the high-type's investment is efficient or not,

$$\begin{aligned} DWL &= (1-f)\sigma_l(C - a_l(\lambda-1)) + f(1-\sigma_h)((\lambda-1)a_h - C) \quad \text{if } \lambda a_h - C > a_h \\ &= (1-f)\sigma_l(C - a_l(\lambda-1)) + f\sigma_h(C - (\lambda-1)a_h) \qquad \text{if } \lambda a_h - C < a_h. \end{aligned} \tag{7}$$

An interior equilibrium could exist where both types of workers should not invest (and indeed would not, if information was complete) but in equilibrium they do.[18]

Note also that an increase in human capital investment does not always improve welfare. I have assumed that the investment of low-ability workers is inefficient, so any investment on

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

their part reduces welfare. Even if the investment of the high-ability worker is efficient, I would still need to weigh the relative loss and gain.

5 Endogenous Cost of College

I 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 nonstandard results in the market. To see the full effects I need to see how the cost of college is determined in equilibrium. I therefore specify how production of education takes place and solve for the equilibrium tuition and quantity of students. I 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

Tuition cost is endogenized by specifying that the production of skill uses scientists who are in limited supply. In particular, assume that a constant fraction of expenditures is spent on these scarce 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 a share of expenditures on research and instruction of around 0.4.[19]

I 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,

$$L_H = S^{\alpha} Y^{1-\alpha}. \tag{8}$$

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

$$\max_{S,Y} CS^{\alpha} Y^{1-\alpha} - wS - p_y Y. \tag{9}$$

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

$$C = \chi \left(w\right)^{\alpha}. \tag{10}$$

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

[19] Data are from The Integrated Postsecondary Education Data System (IPEDS).

The scientists' wage, w, is given by the equilibrium in the scientists' market. From the firms' maximization, the demand for scientists is given by

$$S^d = L_H w^{\alpha-1} (\frac{1-\alpha}{\alpha})^{\alpha-1}. \tag{11}$$

This must be equal to the fixed supply, $\overline{S}$. Substituting for the wage, w, from (10) I have the supply of college given by

$$L_H^s = \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} \left(\frac{C}{\chi}\right)^{\frac{1-\alpha}{\alpha}} \overline{S}. \tag{12}$$

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

I now 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 satisfies 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 the parameters (ϕ, ψ),

$$\frac{f - \psi(C)}{\phi(C) - \psi(C)} = \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} \left(\frac{C}{\chi}\right)^{\frac{1-\alpha}{\alpha}} \overline{S}, \tag{13}$$

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 seen in the 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, I can now look at the comparative statics and show how the model explains the recent trends and what the predictions are 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 rise in tuition fees due to more 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. These explanations differ along the new dimension that the model introduces: self-selection. An $SBTC$ 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 heterogeneous 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

This paper presents an equilibrium model of the college market 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.

The paper contributes to the signaling literature by exploring the possibility that self-selection that is due to differential returns and not differential costs. It shows that an equilibrium with positive selection arises if 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 quality affects prices in an environment of imperfect information. Stating the problem in this way allows me 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 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 implications for inequality and social policy. The first concerns the debate around the ongoing expanded financial assistance for education at state and federal levels. The above 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, Shapiro, and Willen (2005), who use the standard signalling model of Spence (1973), augmented with credit constraints, to reach an opposite conclusion.

The second result compares the education decisions of identifiable groups that differ in their average human capital. The likelihood of going to college increases for all agents in the group who have 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 suggests that a better understanding of the ability composition of skill-groups is needed: 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 tuition costs, and, in turn, how the quality of workers 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. I 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 then 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 the Bayesian beliefs must 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"):

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

Now, express the 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 equations (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 $\frac{\lambda-(A+1)}{\lambda-A}\frac{A}{A+1} > \frac{4ab}{(a+b)^2}$. For the solutions to be probabilities between (0,1) we must further have $0 \le \frac{a(1-b)}{a-b} \le 1$ and $0 \le \frac{(1-b)}{a-b} \le 1$.

(ii) I show that $(\sigma_h, \sigma_l)^* = (1,1)$ can be a part of an equilibrium iff $\frac{p_h p_l}{p_h+p_l} + \frac{(1-p_h)(1-p_l)}{1-p_h+1-p_l} > \frac{A}{\lambda}$. 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 $p_l\lambda\widehat{\phi}_h + (1-p_l)\lambda\widehat{\phi}_l > A \Longleftrightarrow \frac{p_h p_l}{p_h+p_l} + \frac{(1-p_h)(1-p_l)}{1-p_h+1-p_l} > \frac{A}{\lambda}$.

(iii) I 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 a higher probability of a good signal; a_l prefers H if $p_l(-\widehat{\psi}_h) + (1-p_l)(-\widehat{\psi}_l) < A$, $\Leftrightarrow \frac{p_l p_h}{p_h+p_l} + \frac{(1-p_l)(1-p_h)}{(1-p_h)+(1-p_l)} > -A$ (which is always true if $A > 0$).

(iv) I 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 $p_l\frac{p_h}{p_h+\sigma_l p_l} + (1-p_l)\frac{(1-p_h)}{(1-p_h)+\sigma_l(1-p_l)} = \frac{A}{\lambda}$. 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. I 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) I 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): I show that pooling on L fails the divinity 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 such that $\lambda g^i(x^i)=g^i(f)+A$. Because $g^h(f)>g^l(f)$ I 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 for all $x_0\in(x^j,x^i)$ I 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 I 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 *Assumption* 2 *implies* $\phi>f>\psi$ *iff* $\widetilde{p}>\widetilde{f}>\widetilde{q}$.

Proof. of Lemma: By Proportions 2: Assumptions $2\Longrightarrow\phi>\psi\Longleftrightarrow\sigma_h>\sigma_l$. 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) I get $\frac{d\psi}{dC}=-\frac{1}{\lambda(a_h-a_l)p_hp_l(1-p_h)(1-p_l)}\frac{\left(\frac{\lambda p_hp_l}{\widetilde{p}^2}\right)-\left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{p})^2}\right)}{\left(\frac{1}{\widetilde{p}^2}\right)\left(\frac{1}{(1-\widetilde{q})^2}\right)-\left(\frac{1}{(1-\widetilde{p})^2}\right)\left(\frac{1}{\widetilde{q}^2}\right)}$. The denominator is negative by the Lemma. The Nominator is negative since $\frac{p_hp_l}{(1-p_h)(1-p_l)}<\frac{\widetilde{f}^2}{(1-\widetilde{f})^2}$ 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 leaves $\frac{d(\phi-\psi)}{dC}>0\Longleftrightarrow\left(\frac{(1-p_h)(1-p_l)}{(1-\widetilde{q})^2}\right)-\left(\frac{p_hp_l}{\widetilde{q}^2}\right)+\left(\frac{\lambda p_hp_l}{\widetilde{p}^2}\right)-\left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{p})^2}\right)<0$. But by the Lemma this expression is smaller than $\left(\frac{(1-p_h)(1-p_l)}{(1-\widetilde{f})^2}\right)-\left(\frac{p_hp_l}{\widetilde{f}^2}\right)+\left(\frac{\lambda p_hp_l}{\widetilde{f}^2}\right)-\left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{f})^2}\right)=\left(\frac{p_hp_l}{\widetilde{f}^2}-\frac{(1-p_h)(1-p_l)}{(1-\widetilde{f})^2}\right)(\lambda-1)<0$ 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 I only need to check the derivative with respect to λ. This is messy, but it turns out that the same condition is sufficient for $\frac{d(\phi-\psi)}{d\lambda} < 0$. ■

Proof. of Proposition 8 (skill premium): Begin by differentiating equilibrium ϕ and ψ as in the proof of Proposition 6. Adding the appropriate λ we get, $\frac{d(\lambda\phi-\psi)}{dC} > 0 \iff \left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{q})^2}\right) - \left(\frac{\lambda p_h p_l}{\widetilde{q}^2}\right) + \left(\frac{\lambda p_h p_l}{\widetilde{p}^2}\right) - \left(\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{p})^2}\right) < 0$ which is true by Lemma 6 (proceeding the proof of Proposition 6). Note there is no need for condition 1. (ii) is proved similarly. ■

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] Altonji, Joseph G., "The Demand for and Return to Education When Education Outcomes are Uncertain," *Journal of Labor Economics*, Vol. 11(1) Part 1: Essays in Honor of Jacob Mincer. (Jan., 1993), pp. 48-83.

[5] 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

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

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

[8] 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

[9] 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.

[10] 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

[11] 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.

[12] 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

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

[14] 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

[15] 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

[16] 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

[17] 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.

[18] 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.

[19] 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.

[20] 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

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

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

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

[24] Goldin, Claudia and Lawrence F. Katz, "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005," 2007, *National Bureau of Economic Research*, NBER Working Papers: 12984.

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

[26] 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.

[27] 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.

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

[29] Jaeger, David A. and Marianne E Page, "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education," *Review of Economics and Statistics*, November 1996, v. 78(4), pp. 733-40

[30] Juhn, Chinhui, Kevin Murphy M. and Brooks Pierce, "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, June 1993, v. 101(3), pp. 410-42

[31] 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.

[32] 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.

[33] 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

[34] 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

[35] Lange, Fabian and Robert Topel, "The Social Value of Education and Human Capital," in E. Hanushek and F. Welch, (Eds.) *Handbook of the Economics of Education*. 2006 (1) chapter 8. Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

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

[37] 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

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

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

[40] 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

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

[42] Steinberger, Michael, 2005. "Educational Sorting and the Returns to College: 1977-1999." Unpublished manuscript.

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

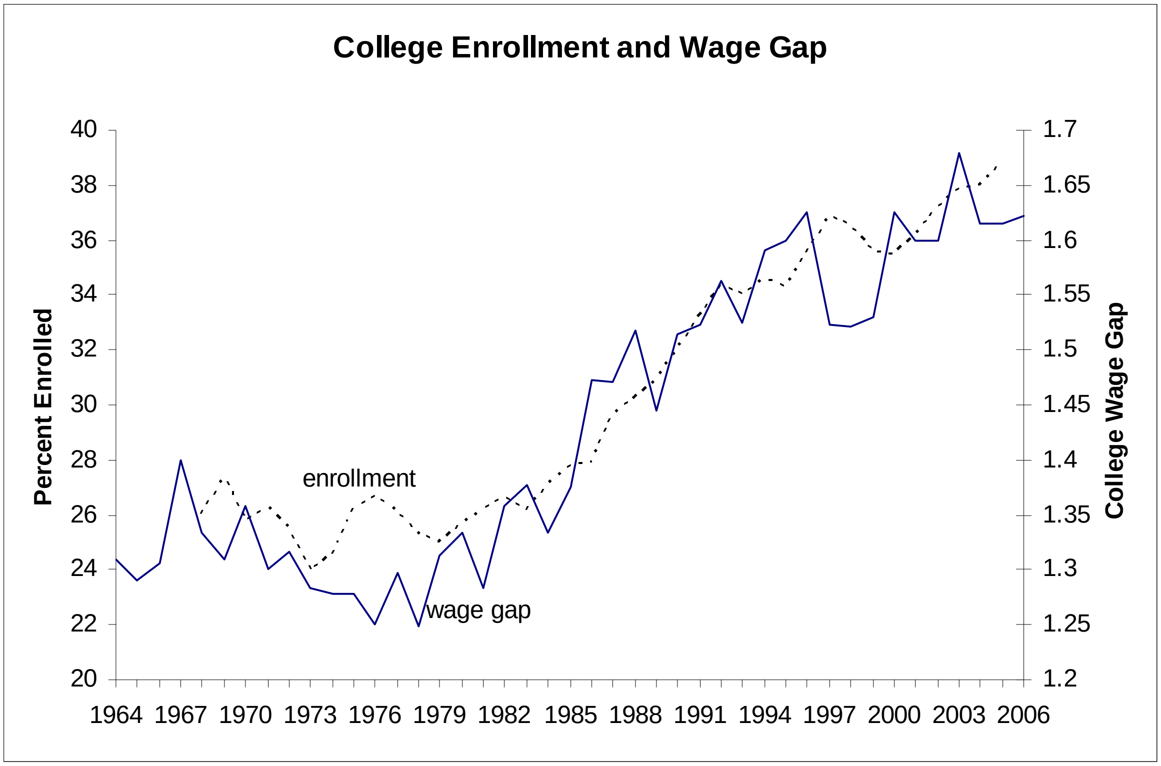


Figure 1:

Source: Current Population Survey (CPS) data for the United States. College enrollement is calculated as the percent of the population 18 to 24 years old enrolled in college. The college wage gap is calculated as the ratio of median earnings of college graduates to high-school graduates. Earnings (in 2006 dollars) are taken for all full-time, full-year wage and salary workers ages 25-34.

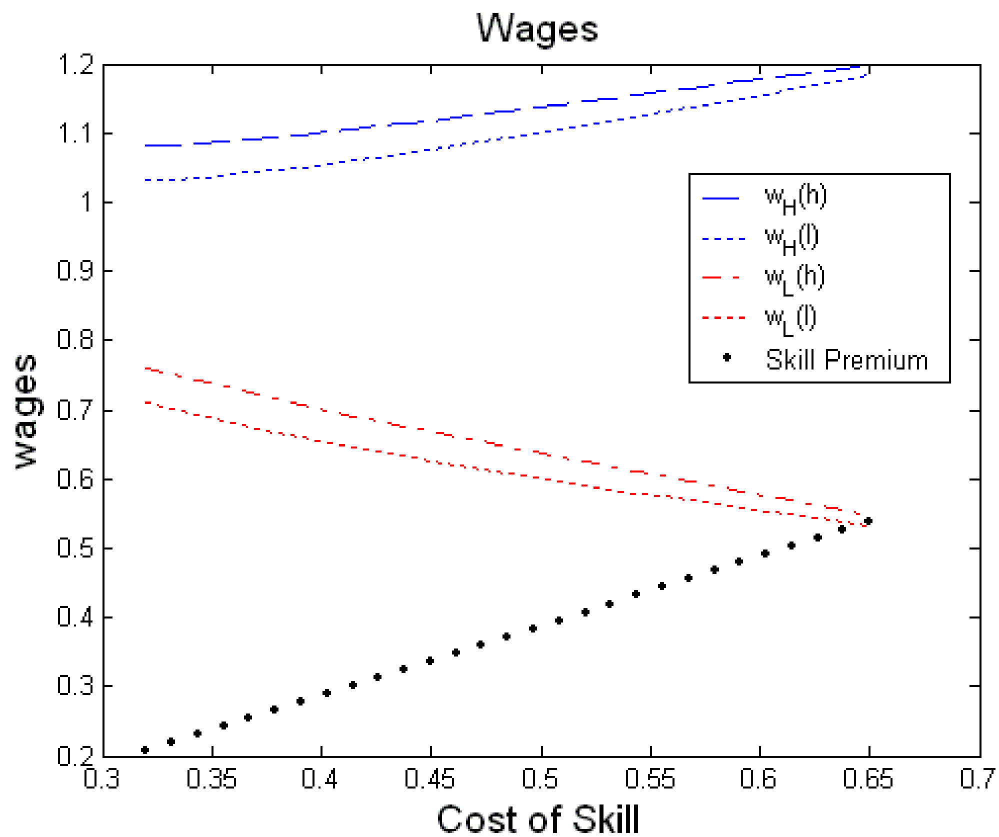
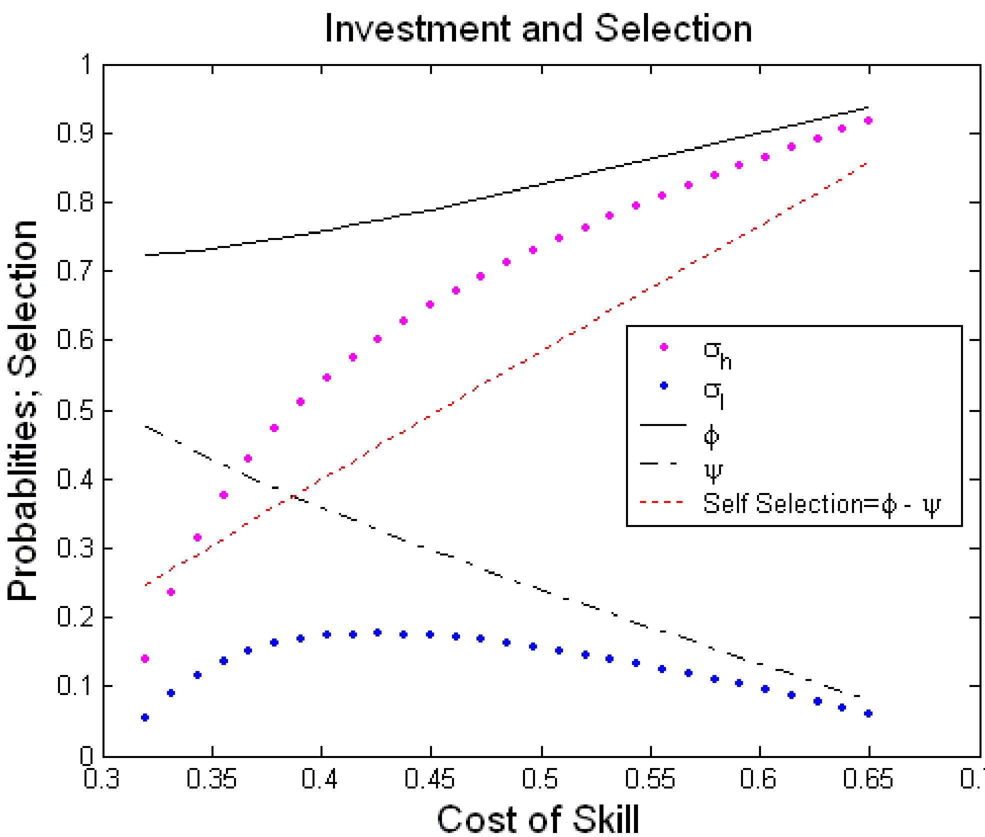

Wages
wages
Cost of Skill
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
$w_H(h)$
$w_H(l)$
$w_L(h)$
$w_L(l)$
Skill Premium

Figure 2:



Investment and Selection
Probablities; Selection
Cost of Skill
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
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 3: