THE PINHAS SAPIR CENTER FOR DEVELOPMENT
TEL AVIV UNIVERSITY

Education Signaling with Uncertain Returns

Tali Regev[1]

Discussion Paper No. 2-12

January 2012

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, an anonymous referee and seminar participants at Bar Ilan, Ben Gurion, Boston, the Federal Reserve Bank of San Francisco, Haifa, Hebrew University, Massachusetts Institute of Technology and Tel Aviv University for their comments. I thank the Pinhas Sapir Center for Development for their financial support and Jeremy Gerst for excellent research assistance.

[1] Tali Regev – The Eitan Berglas School of Economics, Tel Aviv University, Ramat Aviv, Tel Aviv 69978. Email: talir@post.tau.ac.il

Abstract

This paper introduces and explores signaling in the market for education based on imperfectly observed heterogeneity in the returns to education rather than heterogeneity in costs. Workers of heterogeneous abilities face the same costs of education, yet the productivity gain from education is higher for more able workers, and employers can noisily observe productivity. While a separating equilibrium does not exist, the mixed strategy equilibrium is partially revealing: a larger proportion of high ability individuals self select to attend college. The education premium depends on the equilibrium mix of abilities in and out of college, and is increasing in college costs. Adding a production function of college education, 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 and high school graduates.

JEL codes: J24,121,J31

1 Introduction

Since the late 1970 there has been a dramatic increase in the college wage gap while the number of college graduates continued to grow and the cost of college more than doubled (see Figure 1).[1] This paper explores how these trends are related to the quality of college attendees through the role of college education as a signal of workers' ability. Following evidence which suggests this signaling component to education could be sizeable yet untapped, and picking up on the criticism on the monotone inverse relationship between costs and ability required in all signaling models, I develop an alternative signaling model in which heterogenous workers face the same cost of education but expect heterogenous random returns due to employer noisy learning. In equilibrium, the self selection of high ability individuals into college is a continuous measure affected by the cost of college and the demand for skill. The paper presents the necessary and sufficient condition replacing the single crossing property in this context, and introduces some new results, such as human capital externalities. It provides a new signaling model, based on noisy employer learning, of the demand for skills, skill premiums, and college market equilibrium.

This paper departs from the standard signaling framework (Spence 1973) by focusing on uncertain returns to education rather than deterministic costs.[2] Informed workers of heterogeneous abilities face the same costs of education, yet the productivity gain from education is higher for more able workers. Uninformed employers see the education choice of workers and noisily observe productivity using a testing technology. I dispose of the cost based single crossing property assumed in the signaling literature. It is 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 the correlation between ability and costs could even be positive.[3] Here, instead, the productivity gains from college are higher for more able workers.[4] However, the model is not isomorphic to the cost based signaling model, because returns are not deterministically increasing with ability. Rather, returns are expected to increase with ability because of informative testing technology, whereby a worker is more likely to be assessed by his employer as having his true ability than any other ability. Employer learning is thus an essential ingredient for the positive selection of workers into education, and hence for the value of education as a signal. Setting up a signaling model based on employer learning, overturns the insight that the high speed of employer learning limits the value of signalling (Altonji and Pierret 2001, Lange and Topel 2006, and Lange 2007). Here, higher levels of employer learning

[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.

[3] A point made by Griliches (1977).

[4] Altonji (1993) provides evidence that the returns from post-secondary education are higher for more able workers.

does not necessarily hinder the signal value of education, and insterestingly, it may increase inefficient investment in education.[5]

These differences in setup result in a different and richer set of implications. For one, there is no separating equilibrium, and hence the Riley (1979) critique does not apply.[6] The surviving mixed strategy equilibrium can be justified by Harsanyi's (1973) purification argument. This equilibrium is empirically plausible since in reality some workers who do not go to college are more able than some of those who do. Another feature of the equilibrium is that the quality of college attendees—and hence the skill premium—does not depend solely on the exogenous distribution of initial abilities in the population. Rather, individuals' decisions determine the equilibrium mix of abilities of college graduates. I focus on the degree to which the mixed strategy equilibrium is informative, as captured by "self selection" – the quality of college graduates relative to the quality of non-college graduates. Self selection occurs if and only if types would choose separate strategies under full information, a condition replacing the single crossing property in this context.

A novel feature of this signaling model is that the equilibrium self selection is a continuous measure which is affected by parameters of the model, given the information value of the testing technology. Furthermore, the returns from college are endogenously determined, and are increasing with the cost of college and decreasing with the productivity gain from college. Another result is 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. Whenever the shift of low-ability workers away from college increases self selection, the rise in signaling value of education may dominate the initial increase in tuition costs. In such a case the demand for college will slope upwards. The model also predicts human capital externalities: all workers invest more in schooling when the average human capital is initially high. This results because in equilibrium, the quality of each education level is invariant with the initial aggregate level of ability. Hence, when aggregate ability is higher, all workers must increase their investment in education to keep the same equilibrium proportion intact. The implication is that any group with identical observable characteristics, such as race, gender, or nationality, will invest in human capital according to the group's initial ability. This causes a dynamic divergence in human capital between groups with different observable characteristics, and is reminiscent of the statistical discrimination literature.[7]

To shed light on the recent changes in the college market, I close the model by introducing a supply function of college services. College production uses some scientists who are in limited

[5] See also Habermalz (2006), for a model in which high speed of employer learning is not necessarily indicative of a low value of job market signaling.

[6] 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.

[7] Arrow (1973a), Phelps(1972), Coate and Loury (1993). However, the multiple equilibria and coordination failure assumption driving those results is not present here. Closer in spirit is Acemoglu (1996), where increasing returns to human capital stem from a different mechanism, namely, ex-ante investment and costly search.

supply and whose wage is determined in equilibrium. A skill-biased technical change (SBTC), widely believed to have taken place during the period I study, shifts demand toward educated workers.[8] The direct increase in the college premium following an SBTC can be undermined by the lower quality of workers who now choose to become skilled. However, with college enrollment picking up, tuition rises, increasing self-selection, possibly reversing the direct effect of the SBTC. 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.[9]

There is some evidence supporting the education signaling hypothesis. Few 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 to 19 percent net of human capital effects.[10] 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 last 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 of able individuals into higher education.[11] Less definite 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.[12] On the other hand, using direct evidence on pre-college test scores, Murnane, Willett, and Levy (1995) find an increase in the ability composition of college graduates and a decrease in the ability composition of high school graduates. Steinberger (2006)

[8] 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.

[9] 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.

[10] 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.

[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 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.

[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 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.

finds similar results using direct new data on test scores.[13]

Technically, my paper is related to work in information economics dealing with uncertainty in the signaling mechanism. Feldman (2004) and Feldman and Winer (2004) discuss a signaling equilibrium with random costs. Stamland (1999) analyzes a signaling game in which the monotonicity assumption is violated, and discusses the informativeness of the equilibrium. Other related work is Hvid (2003), where education enables learning about ability, however, it is the workers who learn about their true capabilities. Most closely related are Weiss (1983) and Lang and Manove (2011) in which signalling arises from differences in productivity, but under different assumptions than mine.[14] The distinguishing feature of my model is that workers have better knowledge of their abilities than do firms. This is not only a plausible assumption, but also likely to be true for most workers and firms. This essential feature eliminates the possibility of a perfectly separating equilibrium and so, in equilibrium, individuals are mixing strategies so as to get the same net benefit from all education choices. Hence, the model explains why some workers who do not go to college are more able than some of those who do. While this is not the only possible explanation, it is unique in basing this result on workers' better knowledge of their abilities, rather than on unobserved heterogeneity in preferences. Finally my paper is also unique in extending the signaling framework to discuss the equilibrium in the college market.[15]

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 by workers and welfare. Section 5 introduces the college production and solves for the equilibrium in the college market. Section 6 concludes.

2 Model: Signaling with Uncertain Returns

The model describes the signaling decisions workers make based on the private knowledge of their ability type, and the wage offers made by employers. Employer make offers based on the signal and an additional test result indicative of the worker's ability. Workers make their

[13] Murnane, Willett, and Levy (1995) find that the average math score for 1980 male high school seniors who subsequently graduated from college is higher than the average score for the comparable group of 1972 male high school graduates. Steinberger (2006) uses direct new data on pre-college test scores in 1979 and 1999 to find a 4% rise in ability for male graduates with a simultaneous decline in the ability of high school.

[14] Weiss (1983) sets up a signalling model in which ability is not observed by both workers and firms, and a testing mechanism is used for sorting. Contrary to my setup Lang and Manove (2011) assume that the informativeness of the testing technology increases with workers' education level. Also, in their model workers know the returns to education while in my model the returns are stochastic.

[15] Analysis of the college market has either been through the supply of differentiated and competitive college services (Hoxby (1997)), or through the study of demand, and in particular the effect of tuition subsidies on college demand (Feldstein (1995)). Hendel, Shapiro, and Willen (2005) work within a signaling framework, but restrict attention to the demand side, investigating the effect of subsidies on inequality in a signaling equilibrium with credit constraints. As for models of college market equilibrium, Rothschild and White (1995) present a pricing model of college, and Epple et al. (2002) discuss the demographic composition of colleges in market equilibrium.

schooling decision based on their ability type before being tested. Firms offer competitive wages conditional on education and the test outcome, which could be an employment test result or on the job monitoring outcome. I assume for simplicity that workers of various abilities and education levels are perfect substitutes in production. This simplifying assumption makes equilibrium wages depend, in effect, only on the quality of workers with the same observed education level and leaves out the standard quantity effects.[16]

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 $i \in I = \{h, l\}$ by his level of ability $a_i = \{a_h, a_l\}$. A worker's type is private information. The prior distribution of types in the population is common knowledge and is given by $(p(a_h), p(a_l)) = (f, 1 - f)$, so that f is the share of the population with high ability.

A worker of type i chooses his education level $e \in E = \{L, H\}$. A strategy for worker i is a probability distribution over the set of actions, $\sigma_i : E \longrightarrow [0, 1]$. After education is completed, a worker is subjected to an employment test, with an outcome $s \in S = \{h, l\}$ assigning each worker to either the high-ability or low-ability type. The likelihood of a worker i scoring a test outcome s be given by the distribution $p : I \times S \longrightarrow [0, 1]$, which is also common knowledge. For simplicity, test outcomes depend only on the worker's ability and not on his education choice.[17] An important property of the testing technology is that it is informative: the likelihood of a good test score for the high-ability worker is larger than the likelihood of a good test score for a low-ability worker. I will assume throughout this monotone likelihood property:

Assumption 1 *(Monotone Likelihood):* $\frac{p(h,s)}{p(h,s')} > \frac{p(l,s)}{p(l,s')}$ *for* $s > s'$.

Firms observe each worker's education choice e and the noisy test outcome s. A firm which hires l_e^i workers of type i and education 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 educated.[18] A firm's strategy is a wage offer $w_e(s)$ where $w : E \times S \longrightarrow R^+$. Each firm's payoff from hiring a worker with observable education choice and test outcome (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. This is the standard shortcut to replicate a competitive labor market outcome.

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

[17] 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.

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

Worker i's payoff from any pure action e is given by $u_i(e,s,w) = w_e(s) - C(e)$, where $C(e)$ is the cost of acquiring education level e. Note workers of different types pay the same costs of education and if they receive the same test score, they get the same utility. Their payoffs do differ in expectation, $Eu_i(e,s,w) = \sum_s p_i(s) w_e(s) - C(e)$.

Throughout, lower case $\{h,l\}$ denote abilities, while education levels are denoted by upper case $\{H,L\}$. I normalize the cost of acquiring low education to zero, and denote by C the cost of high education, that is, $C \equiv C(H) > C(L) = 0$. Since $\sigma_i(H) + \sigma_i(L) = 1$, I will use $\sigma_i \equiv \sigma_i(H)$ wherever possible to denote the probability of type i going to college. Similarly, since the probabilities for the high type to receive a good or bad test outcome sum up to one, $p(h,h) + p(h,l) = 1$, denote by $p_h \equiv p(h,h)$ the probability of a high type to get a high test outcome, and by $p_l \equiv p(l,h)$ the probability of a low type to get a high test outcome. The monotone likelihood assumption can be therefore stated using this notation as: $\frac{p_l}{1-p_l} < \frac{p_h}{1-p_h}$.

2.2 Equilibrium Definition

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

Denote the expected wage a worker of type a_i gets if he chooses education level e by $Ew_e(a_i) \equiv \sum_s p(s|a_i) w_e^*(s)$. A worker's expected wage from an education choice e depends on his own type through the probability term $p(s|a_i)$, and on the equilibrium composition of his education group through the equilibrium wage term $w_e^*(s)$. Expected wage depend on both individual characteristics and on group composition. The dependency of workers' wages on their ability, reintroduces a type of single-crossing property. At the same time, the information externality is in place because the inferred ability of a worker still depends on the composition of his education group. If in equilibrium the posterior probability of finding a high-ability worker is higher in the educated group than in the uneducated group, then acquiring education has an additional signaling value.

2.3 Types of Equilibria

A worker of ability a_i compares his expected payoff when he acquires education, $Ew_H(a_i) - C$, 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 share of individuals of that type playing the pure strategy education choice. Because I am interested in a non-trivial composition of education groups, 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. *(Firms pay higher wages when observing higher test results).*

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 test result, 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 $Ew_e(a_h) \geq Ew_e(a_l)$.*(A high ability worker expects higher wages than a low ability worker with the same education level).*

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 education level H. By Bayes rule $\mu^*(a_h|H,h) = \mu^*(a_h|H,l) = 1$ so that $w_H(h) = w_H(l) = w_H$. By the previous Lemma, in sector L , $Ew_L(a_l) \leqslant Ew_L(a_h) \leqslant Ew_H(a_h) - C = Ew_H(a_l) - C = w_H - C$ where the second inequality follows from a_h's choice and the equality from the beliefs being $\mu^*(a_h|H,\cdot) = 1$. This contradicts $Ew_L(a_l) > Ew_H(a_l) - C$. A similar argument holds for the a_h fully revealing himself in 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 education choice e, regardless of the test result. But if the test 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 test has power.[19]

[19] 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.

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 education if the social cost of low types investing in education is small enough. On the other hand, pooling on L exists if there is a cost associated with low-ability types investing in education, 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).[20] According to the divinity criterion, an equilibrium can be deleted if there are beliefs regarding off-the-equilibrium-path education choice for which only one type of worker would like to deviate. 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.

[20] 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.

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 education, and the corresponding semi-separating equilibrium has all the high-ability workers investing in education. If it is socially efficient for the low-ability worker to invest in education ($A < 0$), then the corresponding semi-separating equilibrium has all the high-ability workers not getting any education.

I now turn to the fully mixed strategy.

3 Equilibrium Analysis

In the fully mixed strategy equilibrium, each worker type is indifferent between the education choices. To gain more intuition regarding the forces at work, think about the solution in terms of the resulting quality in each education 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.

3.1 A Change of Variables: from Quantities to Qualities

Consider the following change of variables. Let ϕ denote the fraction of high-ability workers in the educated pool, and similarly let ψ be the fraction of high-ability workers in the uneducated 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 education 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 information from the test. 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 strategies σ^* to form their correct beliefs $\mu^*(a_i|e,s)$. Given these beliefs, I can construct expected wages for each education level and test outcome, $w_e^*(s)$. Finally, workers take these wages as given when making their education decision. In the fully mixed strategy equilibrium, the expected returns from a worker's choices must be equalized, $Ew_H^* - C = Ew_L^*$, where expectation are taken over the distribution of test outcomes, a distribution which varies across worker types. This boils down to two equilibrium equations, one for each type of worker i,

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

Since wages $w_e^*(s)$ are based on education and test outcomes and independent of type, but the distribution of test outcomes $p(s|a_i)$ exhibits the monotone likelihood property, these two equations can hold simultaneously only if for each test outcome s, $w_H(s) - w_L(s) = C$. That is, the skill premium must be the same for workers receiving high and low test outcomes, and equal to the cost of acquiring skill, C.

To solve, plug the expressions for wages into these two equations to 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 and simplify the equilibrium equations to,

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

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

Each equation corresponds to a test outcome: equation (3) equates the rewards for a high test outcome ($s = h$) from getting education or remaining uneducated, and equation (4) does the same for a low test outcome ($s = l$). The composition of abilities in each education level, ϕ and ψ, must be such that the returns to each test outcome are the same across education levels. Both conditions impose a positive relationship between ϕ and ψ, the quality of workers in the H and L education levels. This is because the investment decision of high-ability workers has a positive external effect on the value of the education group. To equate the returns across education choices, ϕ and ψ must move together.

These expressions also disentangle the real value of education 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 education. 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 highlights the role of information available to firms. Suppose the testing technology was uninformative, such that high and low ability workers had the same probability of getting a high score, $p_h = p_l$. Then, 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, Human Capital Investment and Welfare

In this section I explore the implication of the equilibrium allocation and wages for the objects of interest: self selection, the skill premium, human capital investment and welfare. I show that self-selection arises if and only if it is inefficient for the low-ability type to invest in education. I provide a sufficient condition for a higher cost of education to increase self-selection, and for a skill-biased technical change to reduce self selection. Next I show that the endogenous skill premium always increases with the cost of education. I then investigate how investment in education responds to price changes in the economy. I show that investment increases in the initial level of human capital. I end with a short discussion of welfare.

4.1 Self-Selection

In any standard signaling model, we say there is self-selection or positive sorting if the equilibrium is separating, that is, if ability increases with education. We need to extend the concept of self-selection to the current setting where there is no separation. As a natural extension, a mixed strategy equilibrium exhibits self-selection if the *average* ability of individuals increases with education. Average ability, in turn, is given by the fraction of high ability individuals within each education level, which we denoted by ϕ and ψ. There is a degree to self-selection, which can be measure by the difference between the qualities in the H and L education levels.

Definition 2 *Let the measure of self-selection be $\phi - \psi$.*

There is self-selection 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 $\sigma_h > \sigma_l$. We care about self-selection and the conditions for it to arise, because only then education acts as a signal of ability. It turns out that a necessary and sufficient condition for self-selection to arise in equilibrium is the following condition, which is hereafter assumed,

Assumption 2 *It is inefficient for the low-ability type to invest in education ($C > a_l(\lambda - 1)$).*

Since C is the cost of education and $a_l(\lambda - 1)$ is the net return from education for a low ability worker under full information, we require that costs exceed benefits from a social planner's perspective. Low ability workers may get private benefits from earning education as their type is not fully known, and their choice of high education exerts a net social cost, $C - a_l(\lambda - 1)$. Since we are only interested in the case in which education serves as a signal for ability, the assumption is crucial, as I now prove:

Proposition 5 *Self-selection arises iff it is inefficient for the lowest type to invest in education. 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 $\sigma_h > \sigma_l$). ■

Intuitively, if under full information the lowest type will not invest in education, than adding any amount of noise will not reverse the higher returns from college for high ability types. Assumption 2 guarantees higher expected returns from education to higher ability individuals, and so serves the same function as the standard single-crossing assumption. Assumption 1 (monotone likelihood assumption) is needed, but is not sufficient for self-selection since it only ensures that information about true types is revealed, and thus that high-ability workers expect higher wages at any education level (see Lemma 3). Note also that I do not need to assume that it is efficient for the high-ability type to invest in education ($a_h < \lambda a_h - C$). There can be a mixed equilibrium with some fraction of both types investing in education even if it inefficient for both of them to invest.

For the next result on the comparative statics of self-selection, I need to impose a condition on the parameters. Define the unconditional probability of getting the good test outcome $\widetilde{f}$,

$$\widetilde{f} \equiv f p_h + (1-f)p_l.$$

Consider the following condition on the parameters:

Condition 1 $\frac{p_h p_l}{(1-p_h)(1-p_l)} < \left(\frac{\widetilde{f}}{1-\widetilde{f}}\right)^2$

Intuitively, the condition requires that a uniform response by high and low ability people moving out of the mixed group, will create benefits in terms of the likelihood of a good test outcome. The condition ties the information value of the test to the ability distribution in the population. Under this condition, the likelihood of getting two good test outcomes from a team of two persons taken from the general population and whose abilities are unknown is higher than the likelihood of getting two good test outcomes from a mixed team of high and low ability. Because ex-ante wages depend on test outcomes, "hiding" in a group of unknown quality is jointly better than getting a test outcome based on true types. To understand this condition better, note that there exists a share f^* of high ability workers such that the inequality holds as an equality. When the share of able workers in the population is higher than f^*, it is beneficial for the mixed team to have employers assume the prior distribution. Note also that this break even share f^* is increasing when the test is not very informative, that is, when p_h is not much larger than p_l.

Proposition 6 *Under condition 1, self-selection increases, $(\phi - \psi) \uparrow$, and the quality of low educated workers declines, $\psi \downarrow$, with:*

(i) increasing costs ($C \uparrow$);

(ii) decreasing skill-bias of technology ($\lambda \downarrow$);

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

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

Proof. See Appendix. ■

The intuition for this proposition is that in a mixed equilibrium the payoffs from both skill choices have to be identical regardless of ability. Therefore any variation in parameters, which reduces the net payoffs from education (holding the composition of types by skill fixed), needs to induce an off-setting increase in the returns from education.

Returns depends on the information inferred from the composition of the educated and uneducated populations given the testing technology. Condition 1 states that the information gain for the educated group from the shift of high and low ability individuals out of education will be higher than the information gain the uneducated group will experience. Under condition 1, a symmetric move by high and low ability persons out of the general population group is increasing the likelihood of a good signal. But then, so does a symmetric move from a self-selected educated group.[21] By requiring that a simultaneous move of high and low ability persons out of education will increase the likelihood of a good signal, Condition 1 ensures that the selectivity of the group will increase.

This condition is more likely to hold when the share of high ability people in the population is high, and when the test is highly informative (as measured by a larger difference between p_h and p_l). In such a case, an increase in the costs of education will reduce investment in education by

[21] since $f < \phi$, condition 1 implies also $\frac{p_h p_l}{(1-p_h)(1-p_l)} < \left(\frac{\phi p_h+(1-\phi)p_l}{1-\phi p_h+(1-\phi)p_l}\right)^2$ since $\frac{p_h p_l}{(1-p_h)(1-p_l)} < \left(\frac{\widetilde{f}}{1-\widetilde{f}}\right)^2 = \left(\frac{f p_h+(1-f)p_l}{1-f p_h+(1-f)p_l}\right)^2 < \left(\frac{\phi p_h+(1-\phi)p_l}{1-\phi p_h+(1-\phi)p_l}\right)^2$.

both ability types, but the disinvestment in education by low-ability workers disproportionately affects the information extraction.

Some interesting comparative statics results are produced under this condition. The ongoing increase in education costs will cause, in this case, an increase in self-selection, that is, the demand for college by quality applicants will increase with tuition costs with a simultaneous decline in the quality of those choosing to stay out of college. In a standard signaling model, in comparison, a sweeping increase in costs will increase the ability of the person just indifferent between going to college and staying out, thus increasing the average ability both in and out of college. Another interesting prediction follows from an increase in the productivity of educated workers, λ, also interpreted as a skill biased technical change. Under condition 1, self selection declines, together with the quality of low-education workers. Thus, the model explains how a skill biased technical change can decrease the quality of low education workers, and hence their wages, a fact which is yet to be understood.

4.2 Skill Premium

In the next subsection I define the skill premium in this model as the difference between the average expected wages in the high and low education levels. I characterize the behavior of the skill premium, and show that it increases with the cost of college. The composition of each education group is an important ingredient of wage determination, and hence of the skill premium. We should thus expect to see an effect of an education group's average ability on the wages each worker earns.

Define the skill premium as the difference between average expected wages of a educated and uneducated workers, where average expected wages for education level e are given by $Ew_e \equiv \sum_{a_i} p(a_i|e) \sum_s p(s|a_i) w_e^*(s) = \sum_{a_i} p(a_i|e) a_i$, that is, taking the average over types and the expectation over test outcomes. The actual realization of the wage is just some noise around these means.

Definition 3 *The skill premium is defined as $Ew_H - Ew_L$.*

Writing out the expressions for expected wages, $Ew_H = \lambda(\phi a_h + (1-\phi)a_l)$ and $Ew_L = \psi a_h + (1-\psi)a_l$, the skill premium is given by

$$Ew_H - Ew_L = (\lambda\phi - \psi)(a_h - a_l) + a_l(\lambda - 1), \tag{5}$$

The skill premium increases with selection, provided that the quality of low education workers does not decline too much.[22] 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.

[22] The condition on the equilibrium outcomes is $d\psi > -\frac{\lambda}{\lambda-1} d(\phi - \psi)$.

However, when we look directly into the effect of parameters on the skill premium (5), we can prove the two following results:

Proposition 7 *(i) Higher investment costs, C, increase the skill premium $Ew_H - Ew_L$.*

(ii) An increase in the real returns to skill, λ, or an increase in the ability gap, $(a_h - a_l)$, directly increases the skill premium but may have 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 increases the premium paid to college graduates. While the parameters' effect on self selection was ambiguous, and depended on condition 1, the wage premium $Ew_H - Ew_L = (\lambda\phi - \psi)(a_h - a_l) + a_l(\lambda - 1)$ depends on the term $\lambda\phi - \psi$. In the expression for the skill premium, the quality of college graduates is multiplied by their productivity λ, a factor which resolves any ambiguity. Costs, which are uncorrelated with abilities or productivity, turn out to have an effect on wages. The intuition is that the returns from both education levels must be equal in a mixed strategy equilibrium. Hence, with higher costs to college, the returns must increase. The quality of college and high school graduates adjusts such that the premium is higher.

The second result suggests that a skill-biased technical change could have created a larger wage dispersion without the endogenous quality adjustments. Under condition 1, an increase in λ decreases self-selection. While educated workers' productivity increases by a higher λ, the firms' willingness to pay for the gain is harmed by the lower quality of the educated.

Note that wages in this economy depend on the composition of the education group through an information externality. Equilibrium wages are constructed using employers' posterior beliefs of a worker's productivity given his education, test outcome, and group composition. Workers benefit from an increase in the quality of their education group regardless of the specific test outcome they eventually get:

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

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

This information externality is due to imperfect information, and is present in all such models. The quality of a worker's education group benefits both low and high ability workers. If a market has such information frictions, prices are set according to the average characteristics of a group. An econometrician observing the data cannot simply control for the composition of the group, since the change in composition also changes market prices. This result highlights the role of composition in determining wages and the skill premium.

4.3 Investment in Human Capital

Think of investment in human capital as the share of individuals choosing education. Looking at the comparative statics for investment, the results are not as definitive. To see this I back out the investment variables σ_l and σ_h, which mechanically decrease with each of the quality variables,

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

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 I take derivatives with respect to cost, productivities, and even the skill productivity parameter 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, education becomes less attractive for both types of workers. However, when the low-ability types retract from school, the quality of educated workers improves. Firms are willing to pay a higher wage for a worker of higher expected ability. This increase in the value of education is due to an increase in its value as a signal on ability. This increase in the relative value of education could potentially dominate the absolute increase in the cost of education. 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 education levels to make the returns (for each test outcome) equal across education levels. This implies some relationship between the quality of workers in each education level regardless of the initial distribution of abilities. Since wages only depend on the endogenous quality, wages too are independent of initial human capital.

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

Proposition 8 *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 (6) 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 education. 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 groups disadvantaged by their initial ability. It means that even high ability individuals within that group reduce their investment in education when the average ability of the group is low. Since any group with identical observable characteristics is subject to a separate market, I can compare groups that differ along the ability dimension. In a more dynamic setting, where investment in education today affects the ability of the next generation, low initial ability of parents lowers their investment in education, which in turn lowers the ability of the next generation. The market is heading toward wage dispersion and increased inequality between groups. To break away from this course of events, early intervention is needed, so as to increase the initial ability as measured, for instance, by the disposition to be productive participants in the labor market. In addition to the direct value added from higher productivity, there will be the additional positive information externality just discussed.

This model can be applied to any demographic group, identified by gender, race or ethnicity. Each such group has its ability distribution and signaling equilibrium. Different demographic groups possibly differ in their pre-college ability distribution, be it because of lacking early childhood education for blacks, low investment in girls' technical skills, or immigrants cultural and linguistic barriers. Ability is thus interpreted as the productivity of an uneducated worker employed in a labor market with some specific technology. As a concrete example, the increase in higher education attainment by females may have been exacerbated by this information externality. The rise in higher education investment by better prepared females has pulled into college females who have not benefited from a similar improvement in their childhood education.

4.4 Welfare

Welfare, too, has an ambiguous response to changes in prices and productivity. This follows directly from the ambiguity of investment. There is a dead weight loss (DWL) associated with 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 education (Assumption 2), only now do I have to specify whether the high-type's investment is efficient or not,

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

Even if it is inefficient for both types of workers to invest in education, there exists an interior equilibrium where they both invest.[23] 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 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 embed the signaling model within a larger context of the college market and consider the general equilibrium consequences of changing market conditions. So 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, the cost of college needs to be determined in equilibrium. This is done by specifying the production side of education and solving 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 a production of education which uses scientists 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.[24]

A higher education sector is added in the following way. To produce L_H college graduates, a general aggregate good Y and some S scientists are used. Scientists are in limited supply, and

[23] This is the same as in the standard signaling environment.
[24] Data are from The Integrated Postsecondary Education Data System (IPEDS).

earn a competitive wage, w. The price of the aggregate good is normalized to one. Production of college graduates, L_H, takes the Cobb-Douglas form with the share of scientists being α,

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

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

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

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

$$C = \chi \left(w\right)^{\alpha}. \tag{11}$$

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

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

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

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

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

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

I now solve for the college market equilibrium.

Definition 4 *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),

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

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 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 can explain the recent trends taking into account the change in selection. Consider first a skill-biased technical change (λ). Under condition 1, selection 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 scenarios 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 the heterogeneity of workers' abilities, and the additional selection effect caused by their choices. 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 a special signaling equilibrium based on uncertain returns rather than deterministic costs, and applies it to the demand for education within an equilibrium model of the college market.

The paper contributes to the signaling literature by exploring the possibility that self-selection is due to uncertain differential returns. It shows that an equilibrium with positive selection arises if it is socially inefficient for low-ability individuals to invest in signaling. Solving the problem in terms of the quality variables instead of the standard quantity variables is useful, since quality affects prices in an environment of imperfect information. Stating the problem in this way allows for clean comparative statics: the wage gap increases with the net cost of investing in the signal.

The paper takes a further step by looking at the general equilibrium implication of the signaling equilibrium for the market for higher education. When the production of college education depends on scientists in limited supply, the equilibrium wages feed back into college tuition, which in turn affect the equilibrium selection. A skill biased technological change may simultaneously increase the quality of college attendees while decreasing the quality of high school graduates. Such selection may help explain not only the growth in high skill wages, but also the decline in the wages of low skill workers observed in the data.

One novel result has important implications for inequality and social policy. An increase in a population's average endowment of human capital increases the education attainment of all individuals in that population. Individuals whose initial human capital is higher, will have larger returns from education. However, also low ability individuals whose human capital remains low will be drawn to earn an education. They now have higher expected wages because there are more able individuals in their reference group. This suggests that early intervention programs which increase potential market productivity have yet another benefit. These programs create a positive externality on agents which 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. Composition does not simply bias our estimation of the "real" skill premium, but rather, it affects wages in a market with incomplete information. It is valuable to understand how the quality of the workforce is endogenously determined by the increased productivity from education and by tuition costs, and, in turn, how the quality of workers affects those same prices.

While the model is framed within an education market, it applies more widely to any signaling environment, where an individual can reveal information about himself through the choices he makes, and there is some independent learning by the uninformed party. The same intuitions applies for instance to insurance markets, where more low risk individuals choose low premium contracts with contingent compensation schemes, or corporate signaling, where managers use devices such as dividends payout to signal the firm's value.[25]

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.

[25] See Rothschild and Stiglitz (1976) for signaling in insurance markets.

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 education 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 $p_L^\epsilon = p(a_h|L) = \frac{\epsilon f}{\epsilon f+(1-\epsilon)(1-f)}$. ■

Proof. of Proposition 2 : Types of Equilibria.

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

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

■

Proof. Now, express the equilibrium in terms of the four posteriors,

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

(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 $0 \leq \frac{a(1-b)}{a-b} \leq 1$ and $0 \leq \frac{(1-b)}{a-b} \leq 1$.

(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 test outcome. 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): 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 5 *A type a_i is deleted for strategy e under the divinity criterion if there is another type a_j such that $D(a_i, e) \cup D^0(a_i, e) \subset D(a_j, e)$.*

I show that pooling on L fails the divinity criterion if both workers prefer education over the equilibrium pooling, when firms believe only able persons acquire education. 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 education over pooling on H, when firms believe only able persons don't acquire education. ■

For the proof of proposition 6 I use the shortcut notation $\widetilde{p}$ for the probability of having a high test outcome conditional on being in the H group, and $\widetilde{q}$ for the probability of having a high test outcome 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: $\phi > \psi \Longleftrightarrow \sigma_h > \sigma_l$. 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 $\frac{p_hp_l}{(1-p_h)(1-p_l)} < \frac{\widetilde{f}^2}{(1-\widetilde{f})^2} < \frac{\widetilde{p}^2}{(1-\widetilde{p})^2}$, where the first inequality hold by assumption and the second inequality by the Lemma. Similarly differentiating for ϕ we get $\frac{d\phi}{dC} = \frac{1}{\lambda(a_h-a_l)p_hp_l(1-p_h)(1-p_l)}\frac{\left(\frac{(1-p_h)(1-p_l)}{(1-\widetilde{q})^2}\right)-\left(\frac{p_hp_l}{\widetilde{q}^2}\right)}{\left(\frac{1}{\widetilde{p}^2}\right)\left(\frac{1}{(1-\widetilde{q})^2}\right)-\left(\frac{1}{(1-\widetilde{p})^2}\right)\left(\frac{1}{\widetilde{q}^2}\right)}$. 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 7 (skill premium): The skill premium $Ew_H - Ew_L = (\lambda\phi-\psi)(a_h-a_l) + a_l(\lambda-1)$ increases if $(\lambda\phi-\psi)$ increases. (i) Begin by differentiating equilibrium ϕ and ψ with respect to C (the expressions are given in the proof of proposition 6). Adding the appropriate λ we get, $\frac{d(\lambda\phi-\psi)}{dC} > 0 \Longleftrightarrow \left[\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{q})^2} - \frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{p})^2}\right] + \left[\frac{\lambda p_hp_l}{\widetilde{p}^2} - \frac{\lambda p_hp_l}{\widetilde{q}^2}\right] < 0$. This inequality holds because both terms are negative: $\frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{q})^2} - \frac{\lambda(1-p_h)(1-p_l)}{(1-\widetilde{p})^2} < 0 \Longleftrightarrow \frac{1}{(1-\widetilde{q})^2} < \frac{1}{(1-\widetilde{p})^2}$ which is true by Lemma 6. Similarly, $\left[\frac{\lambda p_hp_l}{\widetilde{p}^2} - \frac{\lambda p_hp_l}{\widetilde{q}^2}\right] < 0 \Longleftrightarrow \frac{1}{\widetilde{p}^2} < \frac{1}{\widetilde{q}^2}$ which was proved in Lemma 6. Note there is no need for condition 1. (ii) by similarly differentiating, it can be shown that in these cases the sign of the effects is indeterminate. ■

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] Altonji, Joseph G. and Charles R. Pierret, "Employer Learning And Statistical Discrimination," *The Quarterly Journal of Economics*, MIT Press, February 2001, v. 116(1), pp. 313-350.

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

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

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

[9] 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

[10] 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.

[11] 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

[12] 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.

[13] 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

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

[15] 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

[16] 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

[17] 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

[18] 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.

[19] 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.

[20] 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.

[21] 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

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

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

[24] Epple, Dennis, Richard Romano and Holger Sieg, "On the Demographic Composition of Colleges and Universities in Market Equilibrium," *The American Economic Review Papers and Proceedings*, May 2002, Vol. 92(2), pp. 310-314.

[25] Feldman, David, "Separating Signaling Equilibria under Random Relations between Costs and Attributes: Discrete Attributes", *Mathematical Social Sciences*, 2004, v. 48, pp. 93-101.

[26] Feldman, David and Russell S. Winer, "Separating Signaling Equilibria under Random Relations between Costs and Attributes: Continuum of Attributes", *Mathematical Social Sciences*, 2004, vol. 48, pp. 81-91.

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

[28] 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.

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

[30] Harsanyi, John C. "Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points." *Internationsl Journal of Game Theory,* 1973, v. 2(1), pp. 1–23.

[31] 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.

[32] Steffen Habermalz, "The Speed of Employer Learning and Job Market Signaling Revisited," IZA Discussion Papers 2309, 2006, Institute for the Study of Labor (IZA).

[33] 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.

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

[35] 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

[36] 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

[37] 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.

[38] 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.

[39] 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

[40] Lange, Fabian, "The Speed of Employer Learning", *Journal of Labor Economics*, University of Chicago Press, v. 25(1), January 2007, pp. 1-35.

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

[42] Lang, Kevin and Michael Manove, "Education and Labor Market Discrimination," *American Economic Review*, June 2011, vol. 101(4), pp. 1467-96.

[43] 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.

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

[45] 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

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

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

[48] Rothschild, Michael and Joseph Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," The Quarterly Journal of Economics, 1976, v. 90(4), pp. 629-649.

[49] 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

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

[51] Stamland, Tommy, "Partially Informative Signalling," *Journal of Economic Theory*, November 1999, Vol 89(1), pp. 148-161.

[52] Steinberger, Michael, 2006. "Educational Sorting and the Returns to College: 1977-1999." Unpublished manuscript. http://economics-files.pomona.edu/steinberger/research/Steinberger_College_Sorting_Sept2006.pdf

[53] 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

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

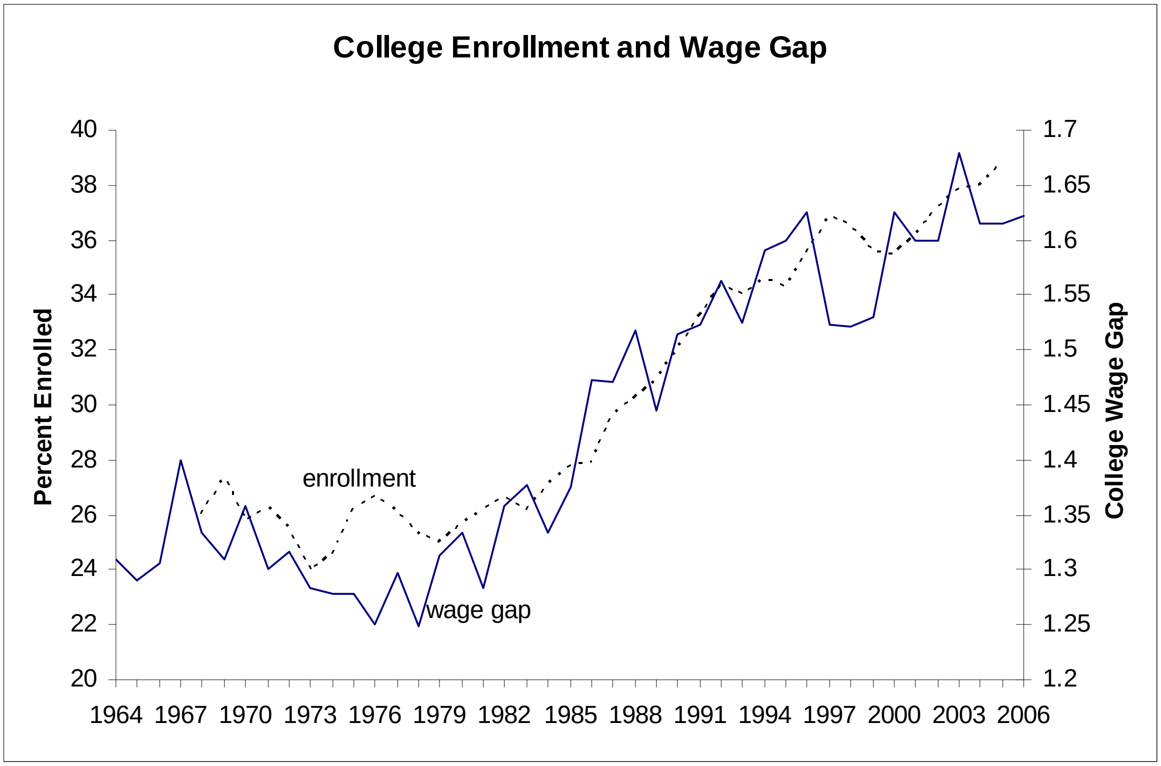

Figure 1:

Source: Current Population Survey (CPS) data for the United States. College enrollment 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.

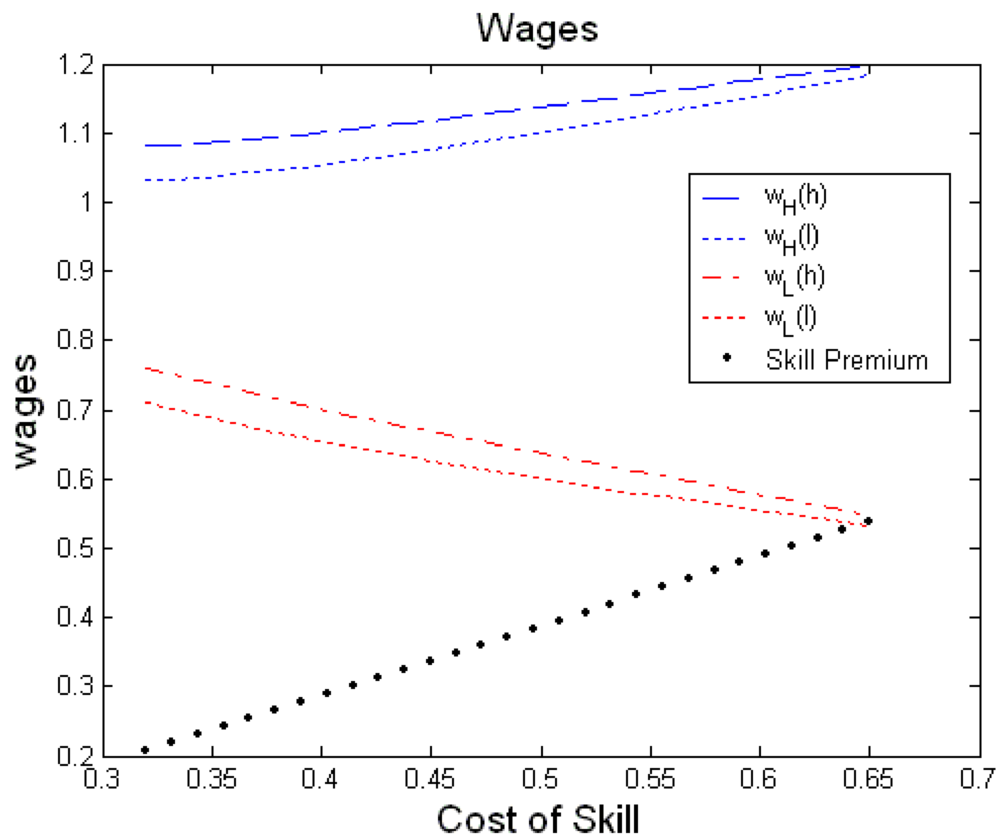
Wages
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
wages
Cost of Skill
$w_H(h)$
$w_H(l)$
$w_L(h)$
$w_L(l)$
Skill Premium

Figure 2:

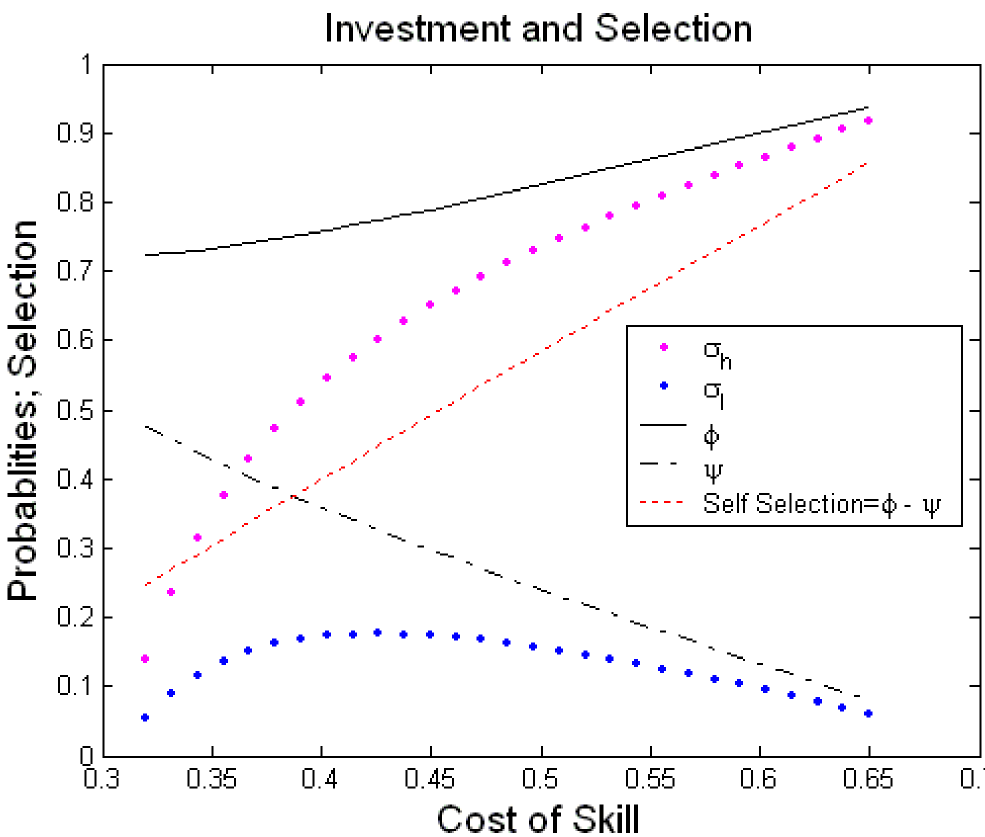
Investment and Selection
Probablities; Selection
Cost of Skill
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
σ_h
σ_l
ϕ
ψ
Self Selection=ϕ - ψ

Figure 3: