Skill Premium, College Enrollment and Education Signals

Yu Zheng
Washington University in St. Louis

June 13th, 2009

Abstract

This paper asks if "higher education as a signal" helps explaining the comovements between college enrollment rate and skill premium for younger workers in the US from 1973 to 2005. In my model a continuum of agents, heterogeneous in talent and initial wealth, make schooling and working decisions: work now or take up college first? When college is very expensive only the wealthy can afford it, hence the degree does not signal much as far as talent is concerned. When college becomes more affordable the degree is a better signal of talent. If talent is valuable, per se, on the work place, the college premium should increase. The model is calibrated to match the basic observed moments of college enrollment and skill premium. For a certain class of production functions it can potentially generate almost all of the growth in college premium after 1985.

Keywords: College Premium, education signalling, college enrollment

JEL codes: J31, D82

*I thank my advisor, Michele Boldrin, for his continuous encouragement and excellent comments. I have also benefitted greatly from discussions with Gaetano Antinolfi, David Levine, Juan Pantano, Raul Santaeulalia-Llopis and Ping Wang. Financial support from Center for Research in Economics and Strategy (CRES), in the Olin Business School, Washington University in St. Louis, is gratefully acknowledged. All errors remain mine. Contact: Seigle Hall 307, Department of Economics, Washington University, St. Louis, MO 63130, U.S. Email: yzhengc@artsci.wustl.edu
1 Introduction

The rise in the college wage premium - defined as the differential between the wages of college and high school graduates - is a well-documented fact. As Card and Lemieux (2001) have shown, the skill premium evolves differently for different age groups: younger workers (age 20 to 40) account for most to the growth of the premium. This paper asks: how much of this evolution can be reasonably explained by the idea that higher education is (also) a signal of talent? The intuition is simple: as college education becomes more affordable the degree becomes an increasingly clear signal of talent; if talent, per se, is useful but unobservable in the working place, the college degree will be rewarded by an increasing premium. The data we seek to understand are reported in the following graph\footnote{For data source and construction, please refer to Section 3.1.}:

![Graph showing College Enrollment Rate and College Premium: U.S. 1969-2005](image)

The evolution of the aggregate skill premium is described, among others, by Autor, Katz and Kearney (2008) and various theories have been put forward to explain it. Katz and Murphy (1992) provide a supply and demand framework to account for the dynamics of wage distribution. Autor, Katz and Krueger (1998) rely on skill-biased technological change to rationalize the demand for skilled labor outpacing the supply. While their model involves assumptions on the unobservable, quality of labor, Krusell, Ohanian,
Rios-Rull and Violante (2000) show that the capital-skill complementarity can account for almost all of the growth in aggregate skill premium without any change in the unobservable. However, Card and DiNardo (2002) question the validity of these theories by pointing out that the skill premium does not grow at the same rate across age groups. Further to this, Card and Lemieux (2001) estimate a production model with imperfect substitution between workers from different age groups and attribute the rising college premium for younger workers to the slowdown of the growth in educational attainment starting with the 1950 cohorts. In contrast to all the aforementioned theories, this paper builds on the idea that education serves as a signaling device and asks how much it can help generating the observed growing college premium. While the application of this signaling story to understanding the trend in wage dispersion is new, the idea of education-as-a-signal is obviously not: it dates back to Spence (1973). Hendel, Shapiro and Willen (2005) bases their theory on similar intuition, but they ask a different question and this is independent work.

In the empirical literature, however, there has been an ample body of research since 1970s, testing the educational screening hypothesis against the traditional human capital accumulation hypothesis, reaching mixed conclusions: Riley (1979), Hungerfor and Solon (1987), Heywood (1994), Lang and Kropp (1986), Bedard (2001) find empirical evidence in support of the signaling story, while Layard and Psacharopoulos (1974), Wolpin (1977), Groot and Oosterbeek (1994) and Albrecht (1981) reach the opposite. Riley (2001) has a summary in its Section 5. Having quantified the two hypotheses within a unified framework, Fang (2006) finds that college education enhances attendees’ productivity by about 40% and productivity enhancement accounts for close to two-thirds of the college wage premium. Taber (2001) offers evidence in support of this paper, in that he finds the increasing premium is more plausibly a result of increasing demand in unobservable ability than in skills acquired in school.

The paper is organized as follows. Section 2 presents the theory, while Section 3 simulates the model and compares its predictions with the data.
Section 4 concludes. All proofs are in the Appendix.

2 Model

2.1 A Static Model: the Working of the Education Signal

A static model may help the reader’s intuition. Assume personal talent is private information that is nevertheless useful in production. Firms can base their wage offer only on the observable signal, which consists of having attained, or not, a college degree. Everyone is born with a high school diploma.

The population has size one, half is endowed with high talent, $\bar{\theta}$, and half with low talent $\theta$. Let the distribution of wealth in the population be $F(\Omega)$. College education has a fixed cost of $Q$. Assume that all those with wealth $\Omega > Q$ go to college, hence, the fraction of people who goes to college is $F(Q)$. Assume there is randomness in successfully completing college. The probability of a high (low) talent person to complete college is $p$ ($p$), with $p > p$. The wage offer is simply the expected talent conditional on the signal received.

With some algebra, we have the wage offer to college graduates,

$$W = \frac{\bar{p}}{p + \bar{p}} \bar{\theta} + \frac{p}{p + \bar{p}} \theta,$$

and to high school graduates,

$$W = \frac{1 - \bar{p}[1 - F(Q)]}{2 - (\bar{p} + p)[1 - F(Q)]} \bar{\theta} + \frac{1 - p[1 - F(Q)]}{2 - (\bar{p} + p)[1 - F(Q)]} \theta.$$

While $W$ is a constant, $W$ depends on the fraction of people that can afford to go to college. Write $x = 1 - F(Q)$, we have

$$W'(x) < 0,$$

implying that the wage differential increases together with college attendance. Next we embed this simple mechanism in a dynamic model of production.
2.2 The Working of the Signal in a Dynastic Model

This is a continuous time discrete-choice problem. Each agent is indexed by the pair $(\theta, k_0)$, where $\theta$ denotes talent, distributed in $[0, \theta]$ according to a cumulative distribution function $G(\theta)$, and $k_0$ is the initial endowment of capital from a distribution $F(k_0)$ over $[0, k_0]$. The distributions $G(\cdot)$ and $F(\cdot)$ are independent. Each agent is endowed with 1 unit of labor. In each instant, an agent faces a discrete choice of whether going to college or not. There are two implicit assumptions in this formulation. One, the offspring of the high (low) type remains high (low); since our main concern is not about social mobility, this assumption seems innocuous. Two, firms cannot, through repeated interaction with an agent from the same dynasty, infer her type. Agents save a constant fraction of their income in each instant. Saving must be positive, i.e. agents cannot borrow against future income. College education requires a fixed cost $Q > 0$. The rest is the same as in the static model, with $p(\theta)$, a monotone increasing function, representing the probability of completing college for type $\theta$.

2.2.1 The Agents’ Problem

In each instant of time, an agent $(\theta, k_0)$ decides whether to go to college or directly to the labor market. If he decides to go to college, he pays the fixed cost $Q$, after which one of the two possible states of nature is realized: he either completes college or not. After finding a job, he works, consumes and saves a fraction $\sigma$ of his income. Agents are risk neutral and maximize the discounted sum of future consumption taking the rental rate of capital $R(t)$ and the wages $\overline{W}(t), \underline{W}(t)$ as given:

$$U(c(t)) = \int_0^{\infty} c(t)e^{-rt}dt.$$

Since there is no disutility from labor, all agents supply 1 unit of labor inelastically. There is no capital depreciation. For ease of exposition, the time argument is suppressed when it does not cause confusion.
Lemma 1 If it is optimal for an agent with talent $\theta$ to go to college at $t$, then it is optimal for any agent who has talent greater than $\theta$ to go to college at $t$ as long as his current capital holding $k \geq Q$.

Intuitively, for an agent with talent $\theta$ attending college is convenient if

$$p(\theta)(\overline{W} - \underline{W}) - RQ$$

is positive. Because $p(\theta)$ is increasing, this implies the result.

2.2.2 Production

In each period the representative firm rents capital from the households and hires workers. I will look at two different classes of production functions. The first class, call it $P1$, is

$$Y(K, L_H, L_C) = [\lambda L_H^p E(\theta \mid HSG) + \nu K^\rho + (1 - \lambda - \nu)L_C^\rho E(\theta \mid CG)]^{1/\rho}, \rho \leq 1,$$

(P1)

where $L_H$ is the number of high school graduates and $L_C$ is the number of college graduates. Here high school graduates and college graduates are perceived as different inputs, i.e. they are assigned different jobs. The productivity of each group is its average talent, by Law of Large Numbers. Implicitly, college education here is productive in the sense that successfully completing college equips the college graduates with a particular set of skills that allow them to undertake a particular task. The elasticity of substitution between two types of labor is the same as their elasticity with capital. In contrast, the second class of production functions only employs aggregate labor and capital as its inputs, that is, skilled and unskilled labor are perfect substitutes:

$$Y(K, L) = A[\alpha K^\rho + \beta (L \cdot E(\theta))^\rho]^{1/\rho}.$$  

(P2)

In both cases, markets are competitive and the high school (or college) graduates will be paid by their marginal product conditional on the signal.
Later, in the calibration section, I will explore the different quantitative implications of the two production functions. The total stock of capital is

$$K(t) = \int_{0}^{K_0} k(t) dF(k_0),$$

and the total labor supply $L(t) = 1, \forall t$. Following the tradition, skilled labor (or, unskilled) and college graduates (or, high school graduates) are used interchangeably.

### 2.2.3 Equilibrium

**Definition 1** Equilibrium without credit markets

An equilibrium without credit markets of this economy is a list $(c(t), k(t), sh(t))$ for each agent $(\theta, k_0)$ and a list of prices $(R(t), \bar{W}(t), \underline{W}(t))$ given initial capital distribution $F(\cdot)$ and distribution of talent $G(\cdot)$, the exogenous positive saving rate $\sigma$ and the production technology, so that

(i) Agents optimally make schooling decision $sh(K(t))$, given $R(t), \bar{W}(t), \underline{W}(t)$;

(ii) Firm maximizes period profit;

(iii) Factor Markets clear.

To provide an analytically convenient environment, we will look at a special class of the equilibrium defined above, the pooling equilibria in which all agents optimally go to college as soon as they can afford it. Before proving the existence of the pooling equilibria, I will prove the monotonicity of the wage differential in enrollment under the proposed strategy profile, which will be useful in the construction of the equilibrium later. Let $x$ be the fraction of agents who go to school and we have $x = 1 - F(Q)$. The theoretical results here are presented mainly for $P1$. An analogous characterization of equilibria with $P2$ is in the Appendix.

**Lemma 2** For $P1$, under the strategy profile that all types of agents go to college as soon as their current capital holdings $k \geq Q$, for high $\rho$ and
low $Q$, $\ln(\overline{W}/\underline{W})$ is increasing in the fraction, $x$, of agents going to college.

To facilitate interpretation, the wage differential has the following formula

$$\frac{\overline{W}}{\underline{W}} = \frac{1 - \lambda - \nu}{\lambda} \left( \frac{L_C}{L_H} \right)^{\rho - 1} \frac{E[\theta|CG]}{E[\theta|HSG]}.$$

An increase in the attendance will unambiguously lead to a higher ratio of expected talents, $\frac{E[\theta|CG]}{E[\theta|HSG]}$, by exactly the same logic as the static model. Imagine $\rho = 1$, then the wage differential will unambiguously go up. However, for $\rho < 1$, the general equilibrium effect kicks in. Since college graduates become more abundant, its marginal productivity decreases relative to that of high school graduates, and this mitigates the effects of the signals. For every $Q$, I can find a $\tilde{\rho} \leq 1$, such that for all $\rho \geq \tilde{\rho}$, this monotonicity property of the wage gap holds. In general, the monotonicity of wage differential rely also on small $Q$ and high $\rho$. Consider a separating equilibrium, in which higher types opt for school and lower types don’t. Suppose that college is very expensive, hence few people can afford it. Then a college degree is more correlated with wealth than with talent and the signal it contains is weak. The marginal productivity of skilled labor is high, hence skilled labor would be receiving a high payment, if identifiable. But holding a college degree is not such a clear signal of talent, as only the rich can afford it. If college enrollment increases while its cost is constant the signal’s quality does not improve as the high cost of attending college implies we are scraping the "bottom of the barrel" among wealthy people. More generally, this is true also when the cost of attending college decreases as long as it is high and the distribution of wealth is not concentrated at high values of wealth. The marginal productivity of skilled labor decreases, though, relative to that of unskilled labor and, as a result, we may have a range in which increasing college attendance brings about a decrease of the wage premium.

**Proposition 1** Under some assumptions, for $Q$ sufficiently small, there exists a pooling equilibrium where all types of agents choose to go to college as soon as $k \geq Q$. 

7
To guarantee that the net benefit of college attendance

$$p(\theta)(\overline{W}(t) - \underline{W}(t)) - R(t)Q$$

is positive for all $t$, $Q$ cannot be too high. A sufficient upper bound, $\hat{Q}$, is the solution (which exists) to

$$p(0)(\overline{W}(0) - \underline{W}(0)) = v(K(0) - \hat{Q})^{\rho-1}\hat{Q}.$$

**Remark 1** The bound of admissible $Q$, $\hat{Q}$, is

(i) increasing in $x_0$;

(ii) increasing in $p(0)$; necessarily $p(0) > 0$;

(iii) increasing in $K(0)$.

The above proposition has nice implications about the trends of the college enrollment rate and of the skill premium, which are summarized below in Corollary 1 and 2.

**Corollary 1** There is a cut-off level of the initial wealth for a given $t, k_0(t)$, so that for all agents whose endowment $k_0 \geq k_0(t)$, they will choose college education at $t$. That is, the college enrollment rate is increasing over time.

Observe that all agents who haven't attended college accumulate capital in exactly the same fashion:

$$\dot{k} = \sigma[R(t)k + \underline{W}(t)].$$

Therefore, $k_0(t)$ satisfies

$$k_0(t) = Q - \int_0^t \dot{k}(s)ds,$$

where the evolution of $k(s)$ follows

$$\dot{k} = \sigma[R(s)k(s) + \underline{W}(s)], \ 0 \leq s \leq t.$$

Obviously, $k_0(t)$ is decreasing over time along the equilibrium path.
Corollary 2  The wage gap is widening over time along the equilibrium path.

Proof  Immediate from Lemma 2 and Proposition 1.

In fact, the equilibrium path is completely characterized in terms of the aggregate capital, $K(t)$, and the cut-off wealth level, $k_0(t)$:

\[
\begin{aligned}
K(t) &= \sigma Y(K(t) - x(t)Q, 1 - x \int pdG, x \int pdG) \\
k_0(t) &= -\sigma[R(t)Q + W(t)]
\end{aligned}
\]  

(1)

where $Y(K, L_H, L_C)$ is given by (P1), $R(t)$ given by (A1), $W(t)$ given by (A2) and

\[
x(t) = 1 - F(k_0(t)),
\]

with $K(0) = \int_{k_0}^{k_0} k_0dF(k_0)$ and $k_0(0) = Q$; and

$k_0(t) \geq 0$.

I will use this dynamic system to simulate the model in Section 3.

For $P2$, I can establish the existence of the pooling equilibrium under even weaker assumptions. First, notice

\[
\frac{W}{\bar{W}} = \frac{E(\theta|CG)}{E(\theta|HSG)}.
\]

Lemma 2 is valid for any $\alpha, \beta, \rho$ and $Q$ :

Lemma 2’  For $P2$, under the strategy profile that all types of agents go to college as soon as their current capital holdings $k \geq Q$, $\ln(W/\bar{W})$ is increasing in the fraction, $x$, of agents going to college.

Proposition 1’  For $P2$, under the assumption that $Q < K(0)$, for $Q$ sufficiently small, there exists a pooling equilibrium where all types of agents choose to go to college as soon as $k \geq Q$.

The two corollaries continue to hold and the dynamic system that characterize the equilibrium path remains valid with modified production technology and prices.
2.2.4 Measuring the Effect of the Signals

The next question is how much this story can account for the growth in the skill premium. This is of course an empirical question, but here I will derive a theoretical bound of the force of the signals. A widely held opinion is that compositional change in the labor force has little effect on the distribution of wage. This exercise addresses this concern theoretically and hopefully sheds some light on the kind of environment that makes the forces of signals strong.

Krusell et al. (2000) shows that the growth rate in the skill premium can be decomposed into three effects, the relative quantity effect, the relative efficiency effect, and the skill-capital-complementarity effect:

\[ g_{\ln \frac{W}{W'}} = (1 - \rho)(g_{h_u} - g_{h_s}) + \rho(g_{\psi} - g_{\psi_u}) + \frac{u}{1 - \lambda}(\lambda - \varepsilon)\left(\frac{K}{s}\right)\left(g_K - g_{h_s} - g_{\psi_s}\right). \]

In our context, for P1, the above composition breaks down to:

\[ g_{\ln \frac{W}{W'}} \simeq (1 - \rho)(g_{h_u} - g_{h_s}) + (g_{\psi} - g_{\psi_u}) \]

The change in the distribution of signals leads to a change in the average talent given a signal, which amounts to a change in the efficiency of skilled labor relative to that of unskilled labor. To maximize the effect of the signals, we must choose the underlying parameters to maximize the relative efficiency effect \( g_{\psi} - g_{\psi_u} \):

\[
\sup_{G_t(\cdot), p_t(\cdot)} \frac{\int_0^\theta \theta p(\theta) dG - \int_0^\theta \theta p(\theta) dG}{(1 - x(t)\int_0^\theta \theta p(\theta) dG)(\int_0^\theta \theta dG - x(t)\int_0^\theta \theta p(\theta) dG)} x(t).
\]

**Remark 2**

1. \( x(t) \) and \( x(t) \) are conveniently taken as given at each \( t \). Though they are endogenous variables, in the calibration I will choose the values for some free parameters to replicate the enrollment rate trend. So we may well take it as exogenous here.

2. We allow \( G_t(\cdot) \) and \( p_t(\cdot) \) to be time-dependent. This maximizes the possible explanatory power of the signals and makes per period
problem exactly the same. From now on, we will suppress the time subscript $t$.

**Proposition 2** The effect of signals is bounded by the negative growth rate of the fraction of people that don’t attend college:

$$\sup(g_{\psi_s} - g_{\psi_u}) = \frac{\dot{x}}{1 - x} = -g_{1-x}.$$  

This result suggests that the signals work most effectively when the education can perfectly sort out the highest talents. Consider the following example in which there are only two talents, 1 or 0.

**Example 1** There is a fraction of $\epsilon$ (close to 0) of people with talent of 1 and the remaining are of talent 0. As a result, $E(\theta) = \epsilon$. Suppose people with high talent can pass the exam almost surely, while people with low talent have the probability of success decreasing overtime in the following fashion:

$$p_t(0) = \frac{1}{1 + x(t)}.$$  

Note that at each instant of time the probability of success is still weakly increasing in the talents. The exam costs nothing. Then, one can verify that

$$\frac{E(\theta|\text{with degree})}{E(\theta|\text{without degree})} \to \frac{1}{1 - x}, \text{as } \epsilon \to 0.$$  

$$g_{\psi_s} - g_{\psi_u} = \frac{d}{dt} \ln \frac{E(\theta|\text{with degree})}{E(\theta|\text{without degree})} \to -g_{1-x}, \text{as } \epsilon \to 0.$$  

Note that in this example, the sorting mechanism becomes more and more efficient overtime, which also contributes to the growth of the skill premium. This verifies that the suggested bound can be achieved in the limit in the example. However, in the setting where the probabilities of success are constant overtime, we would expect in general slower growth in skill premium. The bottom line is that in an economy in which the
distribution of degrees is highly upward skewed, the education signal has a bigger force.

Now we do a simple counterfactual calculation. Taking the college enrollment rate from 1969 to 2005, compute the $g_{1-x}^2$. Then, I take the skill premium in 1969, and let it grow at the maximum theoretical bound $-g_{1-x}$, I get the fictitious wage gap in the dashed line constrained with the real data:

![Figure 2: Real and Fictitious Wage Gap](image)

The signals, theoretically, can account for all of the growth in skill premium. But as will be clear in Section 3, our hands are tied significantly by the specification and parameterization of the model.

### 2.3 Optimality

In the current environment, there are two potential sources of inefficiency: the information problem represented by the private information of talents and the problem of missing credit market. We will investigate the consequences of these two problems one by one. In both cases, the objective of the social planner is to maximize period total output.

---

2Since in the proof of the above proposition, $\dot{x}$ is assumed to be positive. I simply replace any negative growth in the data with zero.
2.3.1 Benchmark One: Complete Information

Assume a social planner observes the individual talents. For $P1$, the social planner simply chooses $\theta^*$ so that all agents with talent above $\theta^*$ are educated at a cost $Q$:

$$
\Gamma(\theta^*) = \max_{\theta^*} \{ \lambda(1 - \int_{\theta^*} \bar{p}dG)^{\rho-1} (\int_{\theta^*} \bar{\theta}dG - \int_{\theta^*} \bar{p}dG) + v[K - (1 - G(\theta^*))Q]^\rho \\
+ (1 - \lambda - v)(\int_{\theta^*} \bar{p}dG)^{\rho-1} \int_{\theta^*} \bar{\theta}p(\theta)dG}^{1/\rho} \}
$$

s.t. $0 \leq \theta^* \leq \bar{\theta}$.

This problem is not a conventional concave problem and the solution is messy. To retain tractability, let $\rho = 1$.

**Proposition 3** Consider $\rho = 1$ with $P1$. In cases in which $2\lambda \geq 1 - v$ holds or both $2\lambda < 1 - v$ and $(1 - 2\lambda - v)\bar{p}(\bar{\theta}) < vQ$ hold, it is optimal not to provide education at all.

If $2\lambda < 1 - v$ and $(1 - 2\lambda - v)\bar{p}(\bar{\theta}) \geq vQ$, the optimal cut-off in talent $\theta^*$ is given by $$ (1 - 2\lambda - v)\theta^*p(\theta^*) = vQ. $$

In cases where production relies more on the unskilled labor than on the skilled labor, or in cases where the opportunity cost of investing in education is high, it may be optimal not to provide education at all. But with incomplete information, there may still exist pooling equilibria defined in section 2.2.3. The individual incentive to self-signal the talent causes both misallocation of factors and a waste of resources. More generally, in all of the pooling equilibria, after some finite length of time, the economy will always over-invest in education, even though it may never reach the optimal amount of skilled labor even in the limit.

With $P2$, the talents are irrelavant since talents are perfectly substitutable and the social planner simply uses all available resources.
Proposition 3’ For $P_2$, the social planner employs all labor and capital and the period output is

$$A[\alpha K^\rho + \beta(E(\theta))^\rho]^{1/\rho}.$$ 

The point is that with complete information, there is no need to invest in education, if education serves purely as a signal, as is in the case with $P_2$.

2.3.2 Benchmark Two: Relaxing Borrowing Constraints

In this section, agents of the same generation are allowed to borrow from each other. Let $b(t)$ be the amount of debt (or credit) that the agent acquires before he receives his income, which has to be paid back at the end of that period.

Definition 2 Equilibrium with within-generation credit markets

An equilibrium of this economy is a list $(c(t), k(t), sh(t), b(t), R(t), W(t), W(t'))$ for each agent $(\theta, k_0)$, given initial capital distribution $F(\cdot)$ and distribution of talent $G(\cdot)$ and the exogenous positive saving rate $\sigma$ and the production technology, so that

(i) Agents optimally choose $sh(K(t))$ and $b(t)$, given $R(t), W(t), W(t')$;
(ii) Firm maximizes period profit;
(iii) Factor markets clear;
(iv) Credit markets clear:

$$R_0 \int_0^{k_0} b(t; \theta, k_0) dG(\theta) dF(k_0) = 0.$$ 

Notice that Lemma 1 still holds. It is easy to construct an equilibrium in which all agents go to college from day 1.

Proposition 4 Under Assumptions 1- 3 and $P_1$, for $Q$ sufficiently small, there exists an equilibrium in which all agents go to college from day 1.

In this equilibrium, the college attendance rate is always 1 and the wage gap remains unchanged

$$\frac{W}{W} = \frac{1 - \lambda - \nu}{\lambda} \left( \frac{\int pdG}{1 - \int pdG} \right)^{\rho-1} \frac{\int \theta pdG}{\int \theta dG} \int \theta pdG.$$ 

14
It is clear from the proof that for all economies that have an equilibrium with borrowing constraints as defined in *Definition 1*, there is also an equilibrium with within generation credit markets as defined in *Definition 2*, in which there is full attendance. Indeed, the equilibrium with within generation credit markets is easier to support in the sense that it can exist for even higher cost of education. Now the evolution of the aggregate capital is described by

$$K'(t) = \sigma[v(K(t) - Q)^\rho + \Pi]^{\frac{1}{3}}.$$  

For the same set of parameters, the equilibrium with within generation credit markets has more skilled labor, less unskilled labor and less capital. Hence, only in an economy where skilled labor is very productive, the relaxing borrowing constraint may bring about more output. More generally, from a social planner’s point of view, relaxing borrowing constraint does not necessarily lead to a Pareto improvement with transfers, since this allows for more competition through unproductive signals. The equilibrium without credit markets converges to the benchmark equilibrium in the limit.

## 3 Calibration

In this section, I calibrate the model and compare the simulated college premium with its data counterpart. I first discuss the source and construction of the data, methodology and the results follow.

### 3.1 Data

The relevant data series are the log wage gap between college graduates and high school graduates, the college enrollment rate and the college completion rate. More detailed description is in the appendix.

*Skill premium.* The point is that the college premium in the model is cohort-based. That is, the theory predicts that for cohorts that are born more recently when the signaling effect of a degree is stronger face higher

$$\Pi = \lambda(1 - \int pdG)^{\rho-1}(\int \theta dG - \int \theta pdG) + (1 - \lambda - v)(\int pdG)^{\rho-1} \int \theta pdG.$$
premium than what earlier cohorts face. I computed the wage series using the CPS March data from 1973 to 2005. The construction process is essentially the same as Autor, Katz and Kearney (2008), with one modification that I computed the wage differential by age groups.

**College enrollment rate.** The college enrollment rate is available from 1960 to 2006 from the American College Testing Program on NCES website. The enrollment rate is obtained by dividing the total number of college enrollment in a given year by the total number of the high school completers, who graduated from high school and completed GED within the preceding 12 months. The current year skill premium and the enrollment rate six years before belong to the same period in the model.

**College completion rate.** Take the number of bachelor’s degrees conferred by degree-granting institutions each year and divide it by the total college enrollment four years before. The degree data are available by year from 1969 to 2006 from NCES. The model counterpart of this statistic is \( \int_0^\varphi p(\theta)dG(\theta) \), the average passing rate of college-goers.

**Initial income distribution in 1973.** I take the wage/salary income distribution of the full-time-fullyear-employed 40-50 years old in 1973 from CPS March. These people were likely to have children around 20-year-old in the same year. CPS sampling weights are used.

**Cost of college.** The cost of college in the model is the tuition, fees, room and board (TFBR) net grants and aids. The TFBR is available from 1976 to 2005 from College Board and the Grants and Aids are available from 1986 to 2006 on selected years. After interpolating the missing observations linearly, the real net cost is almost constant from 1986 to 2006, averaged at 5467 in 2006 dollars.

### 3.2 Calibration

The complication of the simulation lies in the fact that the model involves the unobservable talent. Therefore, instead of asking how much of the skill premium can be explained by the signals, I ask how much signals can matter if the unobservables behave in the most favorable way for me. This leads
to a two-stage strategy, which I will specify next.

### 3.2.1 General Procedure

**1st stage estimation.** Assume that the average talent conditional on having a college degree follows a linear trend.

\[
ht = \frac{\int_0^\infty \theta p_t(\theta)dG}{\int_0^\infty p_t(\theta)dG} = h_0 + \gamma t.
\]

In words, if the college education becomes more efficient in discriminating talents, the skill content associated with a college degree must be higher. This provides another engine of growth in skill premium. Alternatively, one can interpret the trend as representing the human capital accumulation aspect of the education process. Then, \( \gamma \) would be a measure of the technological change in the relative skill of the workers. We will see the estimates of \( \gamma \) is fairly small.

I use a non-linear-least-square model of the evolution of the wage gap for both production functions. For \( P1 \), I normalize \( E\theta = h_0 = 1 \), take \( v = \frac{1}{3} \), and jointly estimate \( \gamma, \lambda \) and \( \rho \). For \( P2 \), I take \( \alpha = \frac{1}{3}, \beta = \frac{2}{3} \), normalize \( h_0 = 1 \), and estimate \( \gamma \) and \( E\theta \). All I take from this stage is the value of \( \gamma \).

**2nd stage calibration.** Take the evolution of \( ht \) from the 1st stage. I calibrate the parameters in such a way that the model replicates the trend of college enrollment rates.

Then, I simulate the model and compare the model prediction of the wage gap with the college premium in the data.

### 3.2.2 Simulation of the Calibrated Model

**P1** The first stage estimation for \( P1 \) yields \( \gamma = 0.4665\% ^4 \). To gain a sense of the magnitude of \( \gamma \), imagine at the initial period, the college degree has no signal value, then by the 33rd period, the average talent of the college graduates will have grown to about 1.15 times the average talent.

In the second stage, I calibrate the model as follows.

---

\(^4\)Significant at 0.05 level.
<table>
<thead>
<tr>
<th>Model</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M^5$</td>
<td>1.9e4</td>
<td>Decision rule</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.4665%</td>
<td>1st stage estimation</td>
</tr>
<tr>
<td>$h_0$</td>
<td>1</td>
<td>Normalization</td>
</tr>
<tr>
<td>$x_0$</td>
<td>0.4657</td>
<td>College enrollment rate in 1973</td>
</tr>
<tr>
<td>$Q$</td>
<td>5467</td>
<td>Real TFRB net aids averaged over 1986 and 2006</td>
</tr>
<tr>
<td>$F(\cdot)$</td>
<td>--</td>
<td>Income distribution in 1973 multiplied by $F^{-1}(1-x_0)$</td>
</tr>
<tr>
<td>$K_0$</td>
<td>6029.4</td>
<td>Mean of $F(\cdot)$</td>
</tr>
<tr>
<td>$v$</td>
<td>1/3</td>
<td>Average capital share of national income in NIPA</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.9</td>
<td>To guarantee the monotonicity of log wage gap</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.3290</td>
<td>To match the initial log wage gap in 1969</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>1.8e-5</td>
<td>To minimize the distance between model enrollment rate and data</td>
</tr>
</tbody>
</table>

$M$ requires some explanation. $M$ scales the productivity of talent to a scale comparable to that of capital, so that in each period the decision rule

$$p(0)(\bar{W} - \underline{W}) - RQ > 0$$

holds, for the existence of the pooling equilibria. The value of $\rho$ implies strong substitutability among the three inputs. Krusell et al. (2000) estimate the elasticity of substitution between unskilled labor and equipment to be 1.67 and that between skilled labor and equipment to be 0.67, which suggests some substitutability between unskilled labor and the combo of skilled labor and capital. In this model, $\rho$ must be high enough to guarantee the monotonicity of the wage gap. With the above parameterization, the model predicts around 80% of the growth in college premium:

Panel 3: Model prediction of college premium for $P1: h_0 = 1, \rho = 0.9$

---

$^5h_t = M(h_0 + \gamma t)$. To guarantee the existence of the pooling equilibrium, I need $p(0)(\bar{W} - \underline{W}) - RQ > 0$. A sufficient condition is that $\pi(\bar{W} - \underline{W}) - RQ > 0$. The scale of $h_t$ guarantees that.
With the modification that $h_0 = 1.3$, I can accommodate a lower $\rho$

Panel 4: Model prediction of college premium for $P1 : h_0 = 1.3, \rho = 0.7$

P2 The model with $P2$ is calibrated as follows:
<table>
<thead>
<tr>
<th>Model</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M$</td>
<td>100</td>
<td>Decision rule</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.3435%</td>
<td>1st stage estimation</td>
</tr>
<tr>
<td>$h_0$</td>
<td>1</td>
<td>Normalization</td>
</tr>
<tr>
<td>$E\theta$</td>
<td>0.9032</td>
<td>1st stage estimation</td>
</tr>
<tr>
<td>$x_0$</td>
<td>0.4657</td>
<td>College enrollment rate in 1973</td>
</tr>
<tr>
<td>$Q$</td>
<td>5467</td>
<td>Real TFRB net aids averaged over 1986 and 2006</td>
</tr>
<tr>
<td>$F(\cdot)$</td>
<td>---</td>
<td>Income distribution in 1973 multiplied by $F^{-1}(1 - x_0)$</td>
</tr>
<tr>
<td>$K_0$</td>
<td>6029.4</td>
<td>Mean of $F(\cdot)$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1/3</td>
<td>NIPA</td>
</tr>
<tr>
<td>$\beta$</td>
<td>2/3</td>
<td>NIPA</td>
</tr>
<tr>
<td>$\rho$</td>
<td>$-1$</td>
<td>Empirical estimate, see Antras (2004)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$8.453e^{-5}$</td>
<td>To minimize the distance between model enrollment rate and data</td>
</tr>
</tbody>
</table>

The resulting simulation yields an even better match with the actual skill premium:

Panel 5: Model prediction of college premium for P3:

If I re-estimate the 1st stage without allowing for trend in unobservable $E(\theta|CG)$. The model still predicts about half of the growth in college premium.

Panel 6: Model prediction of college premium for P3: $E\theta = 0.8451, \gamma = 0$
Discussion  In general, the model does a pretty good job predicting the evolution of the skill premium, both in terms of the magnitude of the growth and the dynamics after 1985. Remarkably, it also captures the hump of the wage differential in late 1990. This suggests that the trend of college completion rates in those years, coupled with the rising enrollment rates match the actual evolution of the wage gap quite well. However, the model fails to catch the swing during the 1973 to 1983 period. This is a period in which the college enrollment kept rising, while the college premium experienced a temporary contraction. The narrowing-down of the wage differential during the 1970s is robust to different measurement specifications and data sources, for example, see Card and DiNardo (2002). Richard Freeman (1976) and Katz and Murphy (1992) suggest that there was an oversupply of educated workers in the 1970s. The bottom line is that the model is able to predict the college premium reasonably well from 1985 to 2005.

At the end of the Appendix, I also estimate the model with a parametric specification of a static talent and degree distribution over the population, which yields less satisfying results.

4 Conclusion and Extension

Though the idea of education as a job market signal is well known, its application to the evolution of wage distribution hasn’t been well articulated in theory. This paper is such an attempt. I have developed a model with agents heterogeneous in initial wealth and talent, who make schooling
decisions. The growth in the college enrollment rate makes a college degree a better signal of high talent. If talent is useful in production, the college degree will be rewarded a higher premium. This brings about a growing wage gap between college graduates and high school graduates. The model is calibrated, with two specifications of production technologies. In both cases, the model has the potential to explain a fair amount of growth of the college premium under reasonable parameterization. Simplistic as it may seem, the theory has a big potential to explain a wider range of phenomena and has a few directions for extension.

4.1 College Premium or Degree Premium

One immediate extension is to extend the two dimensional choice variable to the multidimensional choice of getting bachelor’s, master’s or doctor’s degree. In the similar spirit, the framework can also be easily adapted to explaining the increasingly high premium of attending elite colleges.

Eckstein and Nagypal (2004) argues that the most important group contributing to the increase in the college wage premium is workers with a postgraduate degree. This is consistent with my theory. The increase in the number of Bachelor’s degrees issued will demand even higher degrees to effectively signal one’s talent, which leads to the growing graduate school premium. It is conceivable that with a continuum of choice of levels of education, that varies from low quality colleges to the Ph.D. programs in top universities, the distribution of the education premium to each will tend to fan out over time as the signals work their way through the distribution.

On the other hand, consider the elite universities and colleges. By casual observation, the best schools are becoming more and more accessible to the high talented students, thanks to more effective admission processes and more generous financial aid. As a result, the degree of elite schools must have become more correlated with talent than before. To estimate the premium of the ‘ivy-league’ and observe its evolution over time would be an interesting empirical question.
4.2 College Admission and Financing

Another way of extending the model is, instead of taking the cost of going to college exogenous, modeling the supply side of the college education. The key to the growing enrollment rate is the relaxation of household budget constraint over time through capital accumulation. But in reality there may be other ways that achieve the same effect. One example is the relaxation of the borrowing constraints, as is studied in Hendel, Shapiro and Willen (2001).

Incorporating a sector of college will be a first step toward a general equilibrium approach. Colleges maximize some objective function by choosing costly admission processes. They can either admit students without much screening or undertake costly selection procedure. Colleges can be endowed with reputation such that in the equilibrium some reputedly good colleges choose to be more selective, but will be compensated by higher prices they charge the students. Students in turn will be compensated by the top college premium. The story is more relevant if we can document the growing tuitions of top-notch schools and the growing returns to elite education.

4.3 Optimal Saving Decision

Finally, the model can also be extended toward a full dynamic model, in which agents optimize over consumption and saving. Intuitively, this will help us more, because since the skill premium is growing over time, for subjective discount rate that is not too high, later cohorts will optimally choose to save more, which will allow their children to go to even fancier colleges or allow them to pursue postgraduate degrees, that will further enlarge the associated higher education premium. Combine a full dynamic model with a multiple or even continuum choice of levels of education would certainly make it a more elaborate model, though possibly analytically intractable. One would want to pay the extra cost of computation for more precise quantitative and policy-oriented analysis. After all, the parsimonious model we have here lays out the essential economic intuition just as well.
5 Appendix

5.1 Theoretical Derivation

Lemma 1 If it is optimal for an agent with talent \( \theta \) to go to college at \( t \), then it is optimal for any agent who has talent greater than \( \theta \) to go to college at \( t \) as long as his current capital holding \( k \geq Q \).

Proof The Bellman equations associated with going to college and not going to college, respectively, are:

\[
rv^c(k(t)) = p(\theta)\{(1 - \sigma)[R(t)(k(t) - Q) + \bar{W}(t)] + \frac{dv_i}{dk}\sigma[R(t)(k(t) - Q) + \bar{W}(t)]\} + [1 - p(\theta)][(1 - \sigma)[R(t)(k(t) - Q) + \bar{W}(t)] + \frac{dv_i}{dk}\sigma[R(t)(k(t) - Q) + \bar{W}(t)]].
\]

s.t. \( k(t) \geq Q \)

\[
rv^{nc}(k(t)) = (1 - \sigma)[R(t)k(t) + \bar{W}(t)] + \frac{dv_i}{dk}\sigma[R(t)k(t) + \bar{W}(t)].
\]

Hence, the value function is simply the max of the two:

\[
v_i(k(t)) = \max\{v^c_i(k(t)), v^{nc}_i(k(t))\}.
\]

Given the wage offers and return to capital, since it is optimal for \((k_0, \theta)\) to go to college,

\[
\Delta(k, \theta) \equiv v^c(k) - v^{nc}(k) = (1 - \sigma + \sigma\frac{dv}{dk})[p(\theta)(\bar{W} - \bar{W}) - RQ] > 0
\]

\[
\Rightarrow \Delta(k, \theta') = (1 - \sigma + \sigma\frac{dv(k; \theta', k_0)}{dk})[p(\theta')(\bar{W} - \bar{W}) - RQ] > 0, \forall \theta' > \theta
\]

Hence, independent of the state variable \( k \), \((k_0, \theta')\) would always prefers college as long as going to college is feasible, i.e. \( k \geq Q \). Q.E.D.
Assumption 1

\[ \rho > 1 - \frac{\left( \int_0^\theta \theta dG - \int_0^\varphi \theta dG \right) (1 - F(Q))}{\int_0^\theta \theta dG - (1 - F(Q)) \int_0^\theta \theta dG}. \]

Lemma 2 For \( P_1 \), under the strategy profile that all type of agents go to college as soon as their current capital holdings \( k Q \), and under Assumption 1, the wage gap \( \ln(W/W) \) is increasing in the fraction of agents, \( x \), that go to college.

Proof Under the specified strategy profile, the output and factor prices are

\[ R(t) = v \Upsilon(K(t) - x(t)Q)^{\rho-1}. \quad (A1) \]

\[ W(t) = (1 - \lambda - v) \Upsilon(x(t) \int_0^\theta \theta dG)^{\rho-1} \int_0^\theta \theta dG, \quad (A2) \]

\[ W(t) = \lambda \Upsilon(1 - x(t) \int_0^\theta \theta dG)^{\rho-1} \int_0^\theta \theta dG - x(t) \int_0^\theta \theta dG \]

\[ \frac{1}{1 - x(t) \int_0^\theta \theta dG}, \quad \text{where } \Upsilon = \left\{ \lambda(1 - x(t) \int_0^\theta \theta dG) \int_0^\theta \theta dG - x(t) \int_0^\theta \theta dG + v(K(t) - x(t)Q)^\rho + (1 - \lambda - v)(x \int_0^\theta \theta dG(1 - x(t) \int_0^\theta \theta dG) \right\}^{\frac{1}{\rho}-1}. \]

\[ \frac{d}{dx} \left( \frac{W}{W} \right) = \frac{1}{1 - x(t) \int_0^\theta \theta dG} \left( \frac{\rho - 1}{x} + \frac{\int \theta dG - \int \theta dG}{\int \theta dG - x(t) \int \theta dG} \right) \]

\[ \geq \frac{1}{1 - x(t) \int_0^\theta \theta dG} \left( \frac{\rho - 1}{x_0} + \frac{\int \theta dG - \int \theta dG}{\int \theta dG - x_0 \int \theta dG} \right) \]

\[ \geq 0, \text{ by Assumption 1 and } x_0 = 1 - F(Q). \]

This implies that \( \ln(W/W) \) is increasing in \( x \). Note that \( \forall Q, \text{ Assumption 1 is not empty}. \quad \text{Q.E.D.} \)

Assumption 2

\[ \frac{1 - \lambda - v}{\lambda} \geq \left( \frac{1 - x_0 \int \theta dG}{x_0 \int \theta dG} \right)^{\rho - 2} \frac{\int \theta dG - x_0 \int \theta dG}{x_0 \int \theta dG}. \]
Assumption 2 guarantees \( \bar{W}(t) > \underline{W}(t) \).

Assumption 3

\[ Q < K(0). \]

Proposition 1 Under Assumptions 1-3 and for \( Q \) sufficiently small, there exists a pooling equilibrium where all types of agents choose to go to college as soon as \( k \geq Q \).

Proof The key is to verify that in the suggested equilibrium, all agents optimally make the schooling decision. By Lemma 1, it is sufficient to look at the agent with the lowest talent and make sure he prefers to go to college. Suppose the college attendance is growing over time.

\[
p(0)[\bar{W}(t) - \underline{W}(t)] - R(t)Q = \Upsilon \{ p(0)[(1 - \lambda - v)(x_0 \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \lambda(1 - x) \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \lambda(1 - x) \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \frac{\theta pdG}{pdG} \} - v(K(t) - xQ)^{\rho - 1}Q, \]

where \( \Upsilon \), as is defined in Lemma 2, is positive. Assumption 2 and Lemma 2 implies

\[
(1 - \lambda - v)(x \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \lambda(1 - x) \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \frac{\theta pdG}{pdG} \] is increasing in \( x \). Now

\[
p(0)[\bar{W}(t) - \underline{W}(t)] - R(t)Q \geq \Upsilon \{ p(0)[(1 - \lambda - v)(x_0 \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \lambda(1 - x) \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \lambda(1 - x) \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \frac{\theta pdG}{pdG} \} - v(K(0) - Q)^{\rho - 1}Q \geq 0. \]

\[ \Rightarrow p(0)[(1 - \lambda - v)(x_0 \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \lambda(1 - x_0 \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG} - \lambda(1 - x_0 \int pdG)^{\rho - 1} \frac{\theta pdG}{pdG}] - v(K(0) - Q)^{\rho - 1}Q \equiv \Psi(Q). \]
By Assumption 3

\[
\frac{d\Psi(Q)}{dQ} = [K(0) - Q]^{\rho-2}[K(0) - \rho Q] > 0,
\]

with

\[
\Psi(0) = 0; \quad \lim_{Q \to K(0)} \Psi(Q) = +\infty,
\]

and by Assumption 2,

\[
p(0)[(1 - \lambda - \nu)(x_0 \int pdG)^{\rho-1} \int \theta pdG - \lambda(1 - x_0 \int pdG)^{\rho-1} \int \theta dG - x_0 \int \theta pdG - 1 - x_0 \int pdG] > 0,
\]

then there exists a \( \hat{Q} \) such that

\[
\Psi(\hat{Q}) = p(0)[(1 - \lambda - \nu)(x_0 \int pdG)^{\rho-1} \int \theta pdG - \lambda(1 - x_0 \int pdG)^{\rho-1} \int \theta dG - x_0 \int \theta pdG - 1 - x_0 \int pdG] > 0,
\]

For all \( Q \leq \hat{Q} \),

\[
p(0)[W(t) - W(t)] - R(t)Q > 0, \forall t.
\]

So by Lemma 1, for \( Q \) sufficiently small, all agents want to go to college as soon as they can afford it. Lastly, for all those who are constrained,

\[
k_i = \sigma[R(t)k_i + W(t)] > 0.
\]

This implies that indeed in the equilibrium there will be an increasing fraction of people who can afford education. Q.E.D.

**Corollary 1** There is a cut-off level of the initial wealth for a given \( t \) so that for all agents whose endowment is above the cut-off level, \( k_0(t) \), they will choose college education at \( t \). That is, the college enrollment rate is increasing over time.

**Proof** An agent starts to go to college at time \( t \) that satisfies

\[
k_0^i(t) + \int_0^t k^i(s)ds = Q.
\]
where the evolution of \( k_i \) follows

\[
\dot{k}_i = \sigma[R(t)k_i(t) + W(t)].
\]

At time \( t \) the fraction of agents that goes to college is

\[
1 - F(k_0(t)),
\]

which is increasing in \( t \), since \( k_0(t) \) is decreasing in \( t \). Q.E.D.

**Lemma 2’** For \( P2 \), under the strategy profile that all types of agents go to college as soon as their current capital holdings \( k \geq Q \), \( \ln(\overline{W}/\underline{W}) \) is increasing in the fraction, \( x \), of agents going to college.

**Proof** Let \( a = E(\theta) \). The gross output is

\[
Y = A\{\alpha(K - xQ)^\rho + \beta[L_H E(\theta|HSG) + L_C E(\theta|CG)]^\rho\}^{1/\rho}.
\]

\[
\overline{W} = \Lambda \beta a^{\rho-1} E(\theta|CG);
\]
\[
\underline{W} = \Lambda \beta a^{\rho-1} E(\theta|HSG);
\]
\[
R = \Lambda \alpha (K - xQ)^{\rho-1},
\]

where \( \Lambda = A\{\alpha(K - xQ)^\rho + \beta[L_H E(\theta|HSG) + L_C E(\theta|CG)]^\rho\}^{1/\rho-1} \).

\[
\ln \frac{\overline{W}}{\underline{W}} = \frac{E(\theta|CG)}{E(\theta|HSG)}
\]
\[
= \frac{\int \theta p_dG}{\int \theta p_dG - x \int \theta p_dG} \frac{1 - x \int \theta p_dG}{\int \theta dG - x \int \theta p_dG}
\]

increasing in \( x \). Q.E.D.

**Proposition 1’** For \( P2 \), under the assumption that \( Q < K(0) \), for \( Q \) sufficiently small, there exists a pooling equilibrium where all types of agents choose to go to college as soon as \( k \geq Q \).
Proof

\[ p(0)(\overline{W} - \underline{W}) - RQ \]
\[ = \Lambda\{p(0)\beta a^{\theta-1}[E(\theta|CG) - E(\theta|HSG)] - \alpha(K - x\tilde{Q})^{\theta-1}\tilde{Q}\} \]
\[ \geq \Lambda\{p(0)\beta a^{\theta-1}[E(\theta|CG) - E(\theta|HSG)] - \alpha(K(0) - Q)^{\theta-1}Q\} \].

By the same token, there exists \( \tilde{Q} \), s.t.

\[ p(0)\beta a^{\theta-1}[E(\theta|CG) - E(\theta|HSG)] = \alpha(K(0) - \tilde{Q})^{\theta-1}\tilde{Q}. \]

For all \( Q \leq \tilde{Q} \),

\[ p(0)(\overline{W} - \underline{W}) - RQ \geq 0, \forall t. \]

Moreover, when this is the case, there will be indeed an increasing number of agents going to college. Q.E.D.

**Proposition 2** The effect of signals is bounded by the negative growth rate of fraction of people that don’t attend college:

\[ \sup(g_{\psi_s} - g_{\psi_u}) = \frac{\dot{x}}{1-x} = -g_{1-x}. \]

**Proof** I proceed in three steps.

Step 1: Transformation.

Let

\[ \tilde{p}(\theta) = \tilde{\theta}p(\theta)g(\theta), \]

which necessarily satisfies

\[ \tilde{p}(\theta) \geq 0, 0 \leq \int_0^{\overline{\theta}} \tilde{p}(\theta)d\theta \leq \overline{\theta}. \]

The problem is equivalent to the transformed problem:

\[ \sup \frac{x\overline{\theta} - \int_0^{\overline{\theta}} \tilde{p}\theta d\theta - \int_0^{\overline{\theta}} \tilde{p}d\theta \int_0^{\overline{\theta}} \tilde{p}dG}{\tilde{\theta}(\theta)\int_0^{\overline{\theta}} \tilde{p}dG} \]

\[ s.t. \tilde{p}(\theta) \geq 0, 0 \leq \int_0^{\overline{\theta}} \tilde{p}(\theta)d\theta \leq \overline{\theta}. \]
Let \( \int_{0}^{\overline{\theta}} \theta dG = a \). This problem can be further formulated as a two-step maximization. Given \( a \),

\[
\sup_{\tilde{y}(\theta)} \frac{\int_{0}^{\overline{\theta}} \theta d\theta - a \int_{0}^{\overline{\theta}} \tilde{p} d\theta}{(\overline{\theta} - x \int_{0}^{\overline{\theta}} \tilde{p} d\theta)(\overline{\theta} a - x \int_{0}^{\overline{\theta}} \theta d\theta)}
\]

\[s.t. \tilde{p}(\theta) \geq 0, 0 \leq \int_{0}^{\overline{\theta}} \tilde{p}(\theta) d\theta \leq \overline{\theta}, 0 \leq \int_{0}^{\overline{\theta}} \theta d\theta \leq a\overline{\theta}\]

Then maximize over all possible \( a \).

Step 2: Change of variables.

Let

\[y(\theta) = \int_{0}^{\theta} \tilde{p}(v) dv.\]

Integration by part gives

\[\int_{0}^{\overline{\theta}} \tilde{p} d\theta = \int_{0}^{\overline{\theta}} \tilde{p}(\theta) d\theta = \overline{\theta} y(\overline{\theta}) - \int_{0}^{\overline{\theta}} y(\theta) d\theta.\]

The problem can be rewritten as

\[\sup_{y(\overline{\theta}), \int_{0}^{\overline{\theta}} y(\theta) d\theta} \frac{(\overline{\theta} - a)y(\overline{\theta}) - \int_{0}^{\overline{\theta}} y(\theta) d\theta}{(\overline{\theta} - x y(\overline{\theta}))(\overline{\theta} a - x \int_{0}^{\overline{\theta}} y(\theta) d\theta)}\]

\[s.t. \left\{ \begin{array}{l}
0 \leq y(\overline{\theta}) \leq \overline{\theta}; y'(\theta) \geq 0; \\
\max\{0, \overline{\theta} y(\overline{\theta}) - a\} \leq \int_{0}^{\overline{\theta}} y(\theta) d\theta \leq (\overline{\theta} - a)y(\overline{\theta}). \end{array} \right\}\]

Step 3: Maximization.

Firstly, \( y(\overline{\theta}) \) and \( \int_{0}^{\overline{\theta}} y(\theta) d\theta \) can take values independently. Secondly, the objective is increasing in \( y(\overline{\theta}) \), but decreasing in \( \int_{0}^{\overline{\theta}} y(\theta) d\theta \). But bigger \( y(\overline{\theta}) \) will increase the lowest level that \( \int_{0}^{\overline{\theta}} y(\theta) d\theta \) can take.

If \( y(\overline{\theta}) \leq a \), then the optimal values are \( y(\overline{\theta}) = a \) and \( \int_{0}^{\overline{\theta}} y(\theta) d\theta = 0 \).

If \( y(\overline{\theta}) \geq a \). Then at the optimum, no matter what value \( y(\overline{\theta}) \) takes, \( \int_{0}^{\overline{\theta}} y(\theta) d\theta = \overline{\theta}(y(\overline{\theta}) - a) \). Substituting this relation into the objective function

\[\sup_{y(\overline{\theta})} \frac{\overline{\theta} - y(\overline{\theta})}{(\overline{\theta} - x y(\overline{\theta}))(1 - x)}.\]

30
It is decreasing in \(y(\bar{\theta})\). Hence, at the optimum, \(y(\bar{\theta}) = a\) and \(\int_0^\bar{\theta} y(\theta) d\theta = 0\).

In both cases, the maximum of the objective function is

\[
\sup(g_{\psi_x} - g_{\psi_a}) = \frac{x}{(\bar{\theta} - xa)(1 - x)}.
\]

Now maximize with respect to \(a\),

\[
\sup(g_{\psi_x} - g_{\psi_a}) = \frac{x}{1 - x} = -g_{1-x}, \text{ as } a \to 0.
\]

Q.E.D.

Proposition 3  Consider \(\rho = 1\) with \(P1\). In cases in which \(2\lambda \geq 1 - \nu\) holds or both \(2\lambda < 1 - \nu\) and \((1 - 2\lambda - \nu)\bar{\theta}p(\bar{\theta}) < vQ\) hold, it is optimal not to provide education at all.

If \(2\lambda < (1 - \nu)\) and \((1 - 2\lambda - \nu)\bar{\theta}p(\bar{\theta}) \geq vQ\), the optimal cut-off in talent \(\theta^*\) is given by

\[(1 - 2\lambda - \nu)\theta^* p(\theta^*) = vQ.
\]

Proof  Differentiate the objective function with respect to \(\theta^*\):

\[g(\theta^*)[(2\lambda + \nu - 1)\theta^* p(\theta^*) + vQ].\]

If \(2\lambda \geq 1 - \nu\), maximum is obtained at \(\theta^* = \bar{\theta}\).

Suppose \(2\lambda < 1 - \nu\).

If \((1 - 2\lambda - \nu)\bar{\theta}p(\bar{\theta}) < vQ\), maximum is obtained at \(\theta^* = \bar{\theta}\).

Otherwise, first order necessary condition requires for \(\theta^* \in [0, \bar{\theta}]\),

\[(1 - 2\lambda - \nu)\theta^* p(\theta^*) = vQ.
\]

SOC at \(\theta^*\) gives

\[\lambda - (1 - \lambda - \nu)[p(\theta^*) + \theta^* p'(\theta^*)] < 0.\]
Hence $\Gamma(\theta^*)$ is a local maximum. To check the global optimality, compare it with $\Gamma(0)$ and $\Gamma(\bar{\theta})$.

$$\Gamma(\theta^*) - \Gamma(0) = [(1 - \lambda - v) - \lambda] \int_{\theta^*}^{\bar{\theta}} (\theta p(\theta) - \theta^* p(\theta^*)) dG > 0.$$ 

$$\Gamma(\theta^*) - \Gamma(\bar{\theta}) = [(1 - \lambda - v) - \lambda] \int_{0}^{\bar{\theta}^*} (\theta^* p(\theta^*) - \theta p(\theta)) dG > 0.$$ 

Hence, $\theta^*$ achieves the global maximum. Q.E.D.

**Proposition 3’** For $P2$, the social planner employs all labor and capital and the period output is 

$$A[\alpha K^\rho + \beta(E(\theta))^{\rho}]^{1/\rho}.$$

**Proof** Immediate. Q.E.D.

**Proposition 4** Under Assumptions 1- 3 and for $Q$ sufficiently small, there exists an equilibrium in which all agents go to college from day 1.

**Proof**

The agents’ problem is now as follows.

If an agent goes to college,

$$rv^c(k(t)) = p(\theta) [(1 - \sigma)[R(t)(k(t) - Q) + W(t)] + \frac{dv^i}{dk} \sigma[R(t)(k(t) - Q) + W(t)]]$$

$$+ [1 - p(\theta)] [(1 - \sigma)[R(t)(k(t) - Q) + W(t)] + \frac{dv^i}{dk} \sigma[R(t)(k(t) - Q) + W(t)]]$$

s.t.\(k(t) + b(t) \geq Q.\)

If not,

$$rv^{nc}(k(t)) = (1 - \sigma)[R(t)k(t) + W(t)] + \frac{dv^i}{dk} \sigma[R(t)k(t) + W(t)].$$

The value function is defined by

$$v^i(k(t)) = \max\{v^c_i(k(t)), v^{nc_i}(k(t))\}.$$
By the same logic as in Proposition 1, I need to check whether \( \forall t, \)
\[
v^{ci}(k(t)) - v^{nci}(k(t)) = p(0)[\overline{W}(t) - \underline{W}(t)] - R(t)Q > 0,
\]

Now the factor prices in the proposed equilibrium are
\[
\overline{W}(t) = (1 - \lambda - v)\overline{\Upsilon}(\int \theta dG)^\rho - 2 \int \theta pdG.
\]
\[
\underline{W}(t) = \lambda \overline{\Upsilon}(1 - \int \theta dG)^\rho - 2(\int \theta dG - \int \theta pdG).
\]
\[
R(t) = v\overline{\Upsilon}(K(t) - Q)^{\rho - 1},
\]

where \( \overline{\Upsilon} = \{\lambda(1 - \int \theta dG)^{\rho - 1}(\int \theta dG - \int \theta pdG) + v(K - Q)^\rho + (1 - \lambda - v)(\int \theta pdG)^{\rho - 1}(\int \theta pdG)\}^{\frac{1}{\rho - 1}}. \)

\[
p(0)[\overline{W}(t) - \underline{W}(t)] - R(t)Q
\]
\[
= \overline{\Upsilon}[p(0)[(1 - \lambda - v)(\int \theta pdG)^{\rho - 2}\int \theta pdG - \lambda(1 - \int \theta dG - \int \theta pdG)]
\]
\[
- v(K(t) - Q)^{\rho - 1}Q},
\]

By Assumptions 1-3,
\[
(1 - \lambda - v)(\int \theta pdG)^{\rho - 2}\int \theta pdG - \lambda(1 - \int \theta pdG)^{\rho - 2}(\int \theta dG - \int \theta pdG)
\]
\[
> (1 - \lambda - v)(x_0 \int \theta pdG)^{\rho - 1}\int \theta pdG - \lambda(1 - x_0 \int \theta pdG)^{\rho - 1}\int \theta dG - \int \theta pdG)
\]
\[
> 0.
\]

\( \exists \tilde{Q}^* \) such that
\[
p(0)[(1 - \lambda - v)(\int \theta pdG)^{\rho - 2}\int \theta pdG - \lambda(1 - \int \theta pdG)^{\rho - 2}(\int \theta dG - \int \theta pdG)]
\]
\[
= v(K(0) - \tilde{Q}^*)^{\rho - 1}.
\]

It is readily seen that
\[
\tilde{Q}^* > \hat{Q}.
\]
\( \forall Q < \tilde{Q}^* \),
\[
p(0)[\overline{W}(t) - \underline{W}(t)] - R(t)Q \geq 0.
\]

33
Hence, everyone attends college at all times, while the wage gap remains constant. Q.E.D.

5.2 Calibration

5.2.1 Data

*Skill premium.* The raw data are taken from the CPS March from 1973 to 2005. Only fullyear fulltime workers that have positive wage and schooling are considered. They are grouped by ages. The relevant age group here is those age 23 to 26. The log deflated weekly wage, which is the income from wage and salary divided by weeks worked, is then regressed on dummies of education, geographic region and race, by sexes. The education is the highest education attainment reported, high school dropouts, high school graduates, some college, college graduates or above. The geographical region is grouped in four, Northeast region, Midwest region, South region and West region. For more definitions on the data precession, please refer to Autor, Katz and Kearney (2008). For each sex, the log wage gap is the difference between the prediction for a white college graduate (but with no graduate degree) who lives in the average geographic region and that for a high school graduate counterpart. The log wage gap is the mean of the log wage gaps of the two sexes, weighted by their hours worked. CPS weights are used. I have explored variations of this basic set-up, including the log 10-year-income gap, the log wage gap between a 23 year old college graduate and a 19 year old high school graduate, among others. The results don’t differ much.

*Initial income distribution in 1973.* Annual income from wage and salary are converted into 2006 dollars by CPI index. CPS weights are used. To match the initial enrollment rate, which is 0.4657, I find the 54th percentile in the empirical income distribution and normalize it to be equal to $Q$. That is,

$$F(Q/\xi) = 1 - 0.4657.$$
Further multiply all income in the sample by $\xi$ and this gives the $F(\cdot)$ in the model. $\xi$ can be thought of as the share of income that goes to educational expenses.

*Cost of college.* The real cost of college, computed using the data published in *Trends in College Pricing 2006* and *Trends in Student Aid 2007*, do not show an obvious trend from 1986 to 2006, as the following graph shows. I take $Q$ to be the average over all these years, which is 5467.

![Graph of real cost of college](image)

### 5.2.2 Estimating Time-invariant Talent Distribution

At the end, I briefly summarize the exercise of estimating a time-invariant talent distribution.

The probabilities of passing as a function of talent is taken to be the following form:

$$p(\theta) = 0.20555e^{1.51(\theta/\overline{\theta})}.$$  

The construction of the probability function comes from the empirical distribution of the scores of NAEP-scaled mathematics assessment of the group of 1992 seniors and their educational attainment in 2000 reported in the National Education Longitudinal Study of 1988. The scores serve as a proxy for talent. Combined with the educational attainment of the test-takers, I approximated the function of probabilities. Assume the talents are distributed normally. In the first stage, I estimate the mean and variance
by minimizing the distance between the skill premium in the model and in the data in each year. I use initial values of $(\bar{\theta}, \sigma_\theta) = (10, 4)$ for all cases. The results are

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\bar{\theta}$</th>
<th>$\sigma_\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>38.5293</td>
<td>15.9206</td>
</tr>
<tr>
<td>0.7</td>
<td>51.7109</td>
<td>14.1723</td>
</tr>
<tr>
<td>0.5</td>
<td>63.7709</td>
<td>13.0543</td>
</tr>
<tr>
<td>0.3</td>
<td>80.6028</td>
<td>12.2018</td>
</tr>
<tr>
<td>0.25</td>
<td>85.7526</td>
<td>11.8134</td>
</tr>
</tbody>
</table>

Then I simulate the model for each value of the $\rho$ listed, using the optimal $\bar{\theta}$ and $\sigma_\theta$ obtained in the first stage. The force of the signal is very weak for all $\rho$. I report one of the findings here:

$\rho = 0.3, \alpha = 1/3, \sigma = 8 \times 10^{-6}$.

References


