Macroeconomic Dynamics, **23**, 2019, 1247–1286. Printed in the United States of America.
doi:10.1017/S1365100517000232

THE ROLE OF EDUCATION SIGNALING IN EXPLAINING THE GROWTH OF THE COLLEGE WAGE PREMIUM

YU ZHENG
City University of Hong Kong

This paper incorporates an education signaling mechanism into a dynamic model of production and asks if "higher education as a signal" helps explain the simultaneous increase in the supply and price of skilled relative to unskilled labor in the United States since 1980. The key mechanism is that if college degrees serve as a signal of unobservable talent and talent is productive at the workplace, then improved access to college will enable a higher fraction of the population to signal talent by completing college, resulting in degrees being a better signal about talent and a widening skill premium. When I assess the contribution of signaling in the model calibrated to the US economy from 1980 to 2003, I find that about 10% of the increase in the skill premium can be attributed to the signaling mechanism, after adjusting for the potential decline in the quality of college graduates.

Keywords: College Wage Premium, Education Signaling, Skill-Biased Technical Change

1. INTRODUCTION

The rise in the skill premium—defined as the ratio between the wage of college graduates and the wage of high school graduates, especially among the younger US workers since 1980, is a well-documented fact [Card and DiNardo (2002), Eckstein and Nagypal (2004), Autor et al. (2008)]. The simultaneous rise in the price and supply of skilled labor relative to unskilled labor suggests a demand shifter for skill in the aggregate production function that outpaces the increase in supply to reward skill (or the equivalent a college degree) with a higher price. In the literature, this process is termed the skill-based technical change (or SBTC hereafter) [Acemoglu (1998), Katz and Murphy (1992), Bound and Johnson (1992)]. In search of the empirical content of this theory, Krusell, Ohanian, Rios-Rull and Violante (2000, or KORV hereafter) show that the SBTC can be interpreted as embodied in the fast-growing stock of capital equipment, which is more complementary to skilled than unskilled labor, generating an increasing demand for skill.

I thank Sumru Altug, Michele Boldrin, Ping Wang, an associate editor and two anonymous referees for their helpful comments. Financial support from the City University of Hong Kong Start-up Grant (No. 7200264) is gratefully acknowledged. Address correspondence to: Yu Zheng, Department of Economics and Finance, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Kowloon, Hong Kong; e-mail: yuzheng@um.cityu.edu.hk.

© 2017 Cambridge University Press 1365-1005/17

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

Complementary to the aforementioned demand-side explanations, this paper formalizes and evaluates a supply-side explanation. To the extent that a skilled labor is "produced" by completing college, how a college degree signal is interpreted will also affect the return to college degrees. If one thinks of college as a device that signals talent by awarding degrees to more able students, then clearly conditioning on attending college, a college degree should correlate strongly with talent. But in a society where access to college is restricted to a small group of families, how a college degree should be interpreted and rewarded in the cross section of young workers depends on the nature of the hurdle one must overcome to reach it.

More specifically, imagine a world where agents differ in talent and wealth, and college awards degrees to more talented agents with higher probabilities. When only wealthy families can afford to send their children to college, the young workers without a college degree will mostly consist of those who do not have access to college. In this case, the expected talent reflected in a college degree will depend on the distribution of talent among the wealthy as well as the selectivity of the college sector in awarding degrees, while the expected talent reflected in a high school diploma will be very close to the average talent in the population. As the economy grows, college education becomes more affordable to a wider spectrum of social strata; a higher fraction of college-age youths can signal their talent by attending and completing college. Now the expected talent contained in a college degree reflects the distribution of talent among the growing rich, which may well decline if the increased access draws in less talented students. But more importantly, the expected talent contained in a high school diploma will decline more steeply, since those with only a high school diploma are more likely to be those who fail to get the college degree. Hendricks and Schoellman (2014) show evidence for the United States that the average ability (measured by test scores in high school) of those who obtain only a high school degree deteriorates substantially relative to the average ability of those who complete college from the 1910 to 1960 cohort.

To relate the compositional change to changes in skill prices, in a stationary environment without technical change, the increase in college premium from improved access is driven by a decrease in the expected talent of high school graduates.[1] In a more general environment with SBTC and productivity growth, the increase in college premium is due to the fact that the wage of high school graduates grows less than the average and a lot less than that of college graduates.

I formalize this idea by incorporating an explicit signaling mechanism into a neoclassical aggregate production function with three inputs: capital, skilled, and unskilled labor. By doing so, the efficiency unit of skilled and unskilled labor, usually interpreted as the factor-specific productivity in the SBTC literature, has a direct structural interpretation as the expected talent of college and high school graduates. As more families can afford to prepare their children for college, the supply of skilled labor in this economy grows. At the same time, the signaling mechanism, by the intuition explained above, implies an increase in the differential

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

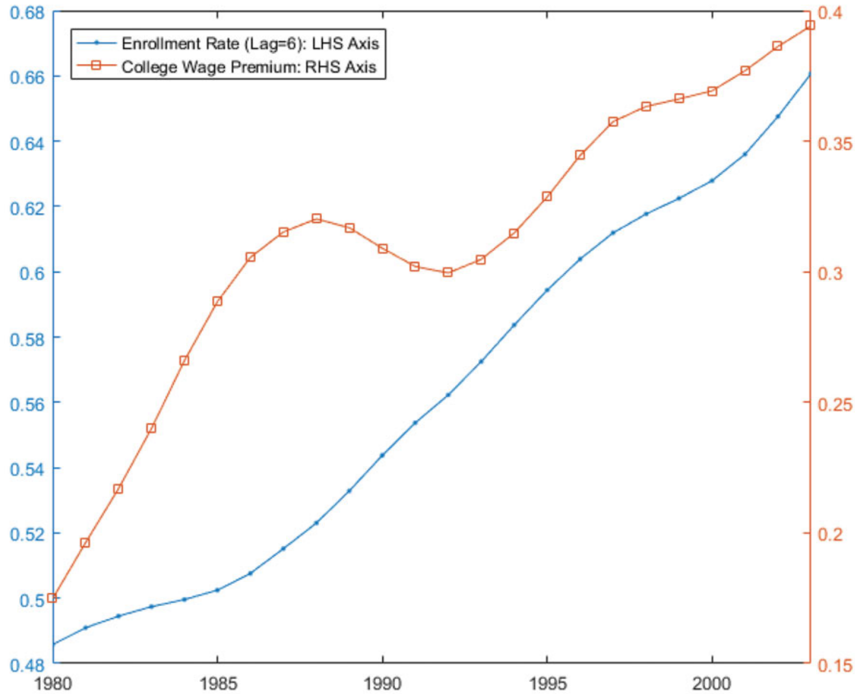

FIGURE 1. HP filtered trend of enrollment rate and college wage premium: 1980–2005. The college wage premium plotted is the log of the ratio of the weekly wage of a college graduate and a high school graduate. The enrollment rates are from *Digest of Education Statistics 2007*. The weekly wage rates are constructed from the CPS March. For the details of the data construction, see Section 3.1.

of the expected talent of college graduates and that of high school graduates, resulting in an upward pressure on the skill premium. The dynamic model has closed-form solutions for the equilibrium path.[2]

I then evaluate the contribution from this signaling mechanism to the growth of skill premium by taking an example of the United States economy from 1980 to 2003. During this period, the United States saw an increasing trend in both college enrollment rates and the college wage premium (Figure 1). The improved access to college in this paper should be interpreted in a broad sense. It can result from an increase in family resources spent on long-term preparation for college [Cameron and Heckman (2001), Carneiro and Heckman (2002)]; the increasing availability of public and private credit to relieve the short-term borrowing constraint [Lochner and Monge-Naranjo (2011), Lochner and Monge-Naranjo (2012)]; an increase in the supply of college [Juhn et al. (2005)]; and the increased affordability of college relative to other consumption goods [Baumol and Blackman (1995)]. The theoretical and quantitative work suggests that the signaling mechanism I model

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

is unlikely to be the sole driver of the skill premium, yet it has a sizeable effect on the US economy, accounting for about 10% of the increase in the college wage premium over the sample period. This estimated contribution from signaling is obtained after a conservative adjustment of the decline of the quality of college graduates as found by Carneiro and Lee (2011).

This paper contributes to the debate on the source of the increase in the college wage premium in the post-1980 US economy. Apart from the aforementioned demand-side explanations, Card and Lemieux (2001), within a supply and demand framework, attribute the increase in the college premium for younger workers to a decrease in the supply of skill from younger workers relative to older workers as a result of the slow down of the educational attainment of younger cohorts. In a general equilibrium overlapping generations model with capital-skill complementarity, He (2012) shows that investment-specific technological change, which shifts the relative demand for skill, is a more important factor in driving the widening skill premium in the postwar US than demographic change, which affects the relative supply of skill [see also He and Liu (2008)]. Based on a human capital interpretation of higher education, Guvenen and Kuruscu (2012) examine the human capital accumulation decision and its implication on the skill premium in a model with SBTC and heterogeneous agents who differ in the ability to accumulate human capital. In comparison, Maoz and Moav (2004) evaluate how endogenous decisions of human capital and physical capital accumulation in a dynamic general equilibrium model with capital-skill complementarity and heterogeneous agents who differ in wealth and access to credit affect patterns of the skill premium. I contribute to the discussion by formally introducing the signaling aspect of education into a model with SBTC and show that, net of the quantity effects, the change of the content of degree signals can also contribute to the rise of the college premium.

The paper is organized as follows. Section 2 contains the theory: Section 2.1 introduces a static model to highlight the basic intuition; Section 2.2 builds a dynamic model of production with education signals; Section 2.3 presents a theoretical upper bound and lower bound on the signaling effect; and Section 2.4 discusses alternative interpretations of talent sorting patterns following improved access to college. In Section 3, I evaluate this model quantitatively for the US economy from 1981 to 2003. The conclusion follows.

2. THEORETICAL FRAMEWORK

2.1. A Static Model: The Working of the Education Signal

A static model may help the reader's intuition. Assume that 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. Initially, in this model, everyone holds a high school diploma. The population has size 1, half is endowed with high talent, $\overline{\theta}$, and half with low talent, $\underline{\theta}$. Let the distribution of wealth in the population be $F(k)$. College education has

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

a fixed cost of Q. Assume that all those with wealth k greater than the cost Q go to college; hence, the fraction of people who goes to college is $1 - F(Q)$. Assume that there is randomness in successfully completing college. The probability of a high (low) talent person completing college is $\overline{p}$ ($\underline{p}$), with $\overline{p} > \underline{p}$. The wage offer is simply the expected talent conditional on the signal received.

With some algebra, the wage offer to college graduates, $\overline{W}$, and to high school graduates, $\underline{W}$, are

$$\overline{W} = \frac{\overline{p}}{\overline{p} + \underline{p}}\overline{\theta} + \frac{\underline{p}}{\overline{p} + \underline{p}}\underline{\theta}. \tag{1}$$

$$\underline{W} = \frac{1 - \overline{p}(1 - F(Q))}{2 - (\overline{p} + \underline{p})(1 - F(Q))}\overline{\theta} + \frac{1 - \underline{p}(1 - F(Q))}{2 - (\overline{p} + \underline{p})(1 - F(Q))}\underline{\theta}. \tag{2}$$

The wage to college graduates, $\overline{W}$, is an average of the high and low talent weighted by the talent-specific probabilities of completing college. The wage to high school graduates, $\underline{W}$, is also a weighted average of the high and the low talent. This weighted average reflects the composition of the pool of high school graduates, which consists of the high talent who attends and fails college, the low talent who attends and fails college, as well as all those who cannot afford college. As a result, the wage to high school graduates depends on the fraction of people that can afford college. Denote the college enrollment by $x = 1 - F(Q)$. It follows that $\underline{W}'(x) < 0$, implying that the wage differential increases together with college attendance. Next I embed this simple mechanism in a dynamic model of production.

2.2. Embedding the Education Signal in a Dynamic Model

This is a continuous-time, discrete-choice problem. The economy is populated by a unit measure of dynastic families. Each dynasty is characterized by the pair (θ, k_0), where θ is the time-invariant talent, distributed over $[0, \overline{\theta}]$ according to a cumulative distribution function $G(\theta)$ in the population, and k_0 is the endowment of capital at time 0, distributed over $[0, \overline{k}_0]$ according to a cumulative distribution $F(k_0)$ in the population.[3] The distributions of initial endowments $G(\cdot)$ and $F(\cdot)$ are independent.[4] An agent at time t is indexed by the dynasty that he belongs to, (θ, k_0). Each agent is endowed with 1 unit of labor, is risk neutral, and maximizes the discounted sum of future consumption in his dynasty, given by the dynastic utility function $U(t; \theta, k_0)$:[5]

$$U(t; \theta, k_0) = \int_t^\infty c(\tau; \theta, k_0)e^{-r\tau}d\tau, \tag{3}$$

where $c(\tau; \theta, k_0)$ is the consumption attained at time τ by the agent, who belongs to the dynasty (θ, k_0), and r is the discount rate.

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

At each instant, an agent faces a discrete choice of whether to go to college or not. If he chooses to go to college, he pays a fixed cost Q, after which he either completes college or not. The probability of completing college is higher for agents with higher talent. This is summarized in the probability of college completion $p(\theta)$, which satisfies $p'(\theta) > 0$ and $p(\theta) > 0$, for all $\theta \in [0, \overline{\theta}]$. Depending on the educational outcome, the agent goes to the labor market either as a college graduate or as a high school graduate. All those who choose not to go to college enter the labor market as high school graduates. Since there is no disutility from labor, all agents supply 1 unit of labor inelastically. There is no capital depreciation. They work, earn wages, receive rental income from the capital, and consume and make an intergenerational transfer, which is a constant fraction ϕ of their income, to the next agent in the dynastic family. I consider only positive transfers. In other words, intergenerational transfer can be made from the older generation to the younger generation only.[6]

Competitive firms hire workers and rent capital for production at each instant. A worker's talent is productive. It is modeled as the efficiency unit per unit labor supply in the production function. Firms, however, do not observe talent, but only observe the educational outcome. Hence, the wage offered to all college (or high school) graduates is the same and reflects the average talent that firms believe the college (or high school) graduates have at the time. Following the tradition, skilled (or unskilled) labor and college (or high school) graduates are used interchangeably.

The agents' problem. At each instant of time t, an agent from dynasty (θ, k_0), or equivalently the representative of the cohort t of that dynasty, makes the schooling decision to maximize the discounted sum of the future consumption of his dynasty, taking the rental rate of capital $R(t)$, the wage of skilled labor $\overline{W}(t)$, and the wage of unskilled labor $\underline{W}(t)$ as given. Write the choice-specific value functions as $v^c(k(t); \theta, k_0)$ for college-goers and $v^{nc}(k(t); \theta, k_0)$ for non-college-goers. Write the value function, taking into account the college attendance decision, as $v(k(t); \theta, k_0)$.

The value of going to college, subject to feasibility given by the budget constraint, satisfies

$$\begin{aligned} rv^c(k(t); \theta, k_0) = p(\theta) \Big\{ &(1-\phi)[R(t)(k(t)-Q) + \overline{W}(t)] \\ &+ \frac{dv}{dk}\phi[R(t)(k(t)-Q) + \overline{W}(t)] \Big\} \\ &+ (1-p(\theta)) \Big\{ (1-\phi)[R(t)(k(t)-Q) + \underline{W}(t)] \\ &+ \frac{dv}{dk}\phi[R(t)(k(t)-Q) + \underline{W}(t)] \Big\} \\ &\text{subject to } k(t) \geq Q. \end{aligned} \tag{4}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

The flow value, rv^c, is equal to the expectation of the sum of the instantaneous payoff, which is given by the current cohort's consumption, and the dynastic capital gain, which is given by the change in the value to the next cohort. The expectation is taken with respect to the probability of completing college. With probability $p(\theta)$, the college-goer completes college, earns the labor income as a skilled worker, and the capital income from his capital holdings (net the cost of college), of which he consumes a fraction $1-\phi$ and leaves the complementary fraction to the next cohort in the dynasty. With probability $1-p(\theta)$, he attends but fails college, earns the labor income as an unskilled worker and the capital income after deducting the cost of college, and splits the total income between own consumption and bequest, according to the rule ϕ.

The value of not attending college satisfies

$$rv^{nc}(k(t);\theta,k_0) = (1-\phi)[R(t)k(t)+\underline{W}(t)] + \frac{dv}{dk}\phi[R(t)k(t)+\underline{W}(t)], \quad \textbf{(5)}$$

where the instantaneous payoff is the agent's consumption, or a fraction $1-\phi$ of the sum of his labor income as an unskilled worker and his capital income, and the dynastic capital gain given by $\frac{dv}{dk}$ times the bequest.

The value function of a cohort-t agent, then, is simply the maximum of the choice-specific values:

$$v(k(t);\theta,k_0) = \max\{v^c(k(t);\theta,k_0), v^{nc}(k(t);\theta,k_0)\}. \quad \textbf{(6)}$$

For ease of exposition, the time argument is suppressed when it does not cause confusion. All proofs are collected in Appendix. I have made the following two observations.

LEMMA 1. *If it is optimal for an agent with talent θ to go to college at t, then it is optimal for any agent who has talent greater than θ to go to college at t as long as going to college is feasible for him:*

$$k(t;\theta,k_0) \geq Q. \quad \textbf{(7)}$$

Intuitively, for an agent with talent θ, attending college is convenient if the net benefit from attending college

$$p(\theta)(\overline{W}-\underline{W}) - RQ \quad \textbf{(8)}$$

is positive. The net benefit from attending college is the expected skill premium minus the cost of a college education. Lemma 1 follows immediately from the assumption that the probability of completing college, $p(\theta)$, is increasing in talent, θ.

LEMMA 2. *If an agent from a dynasty with initial capital endowment k_0 can afford college at t, agents from this dynasty can always afford college at t' greater than t. The fraction of agents who can afford college is increasing in t.*

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

Under a positive rate of intergenerational transfer and zero capital depreciation, the rate of capital accumulation is always positive for all dynasties. Lemma 2 follows naturally.

Production. The production is described by a standard neoclassical production function. It has three inputs: capital k, unskilled labor u, and skilled labor s. I consider a nested Constant Elasticity of Substitution (CES) production function to allow for different elasticities of substitution between the inputs:[7]

$$Y(k, u, s) = A\left\{\mu k^{\sigma} + (1-\mu)[\lambda u^{\rho} + (1-\lambda)s^{\rho}]^{\frac{\sigma}{\rho}}\right\}^{(1/\sigma)} \qquad \sigma, \rho \leq 1, \quad \textbf{(9)}$$

where $u = \psi_u h_u$ and $s = \psi_s h_s$ with ψ_u (ψ_s) being the efficiency unit of unskilled (skilled) labor and h_u (h_s) being the quantity of raw unskilled (skilled) labor input. In our setting, the quantity h_u (h_s) is taken to be the fraction of high school (college) graduates and the efficiency unit ψ_u (ψ_s) the expected talent of high school (college) graduates. When the parameter ρ that governs the elasticity of substitution between skilled and unskilled labor is 1, the two types of labor are perfect substitutes. In the quantitative assessment of this model for the US economy later on, college serves as both a signaling device and a venue for human capital production. In that case, high school graduates are only an imperfect substitute for college graduates, who are equipped with the special know-how acquired in college. This can be captured by an elasticity of substitution parameter ρ that is in between 0 and 1, in which case the actual elasticity of substitution, $\frac{1}{1-\rho}$, is in between 1 and infinity.

The skill premium π, or the ratio between the wages of skilled and unskilled labor, has the familiar form

$$\pi = \frac{1-\lambda}{\lambda}\left(\frac{h_u}{h_s}\right)^{1-\rho}\left(\frac{\psi_s}{\psi_u}\right)^{\rho}. \qquad \textbf{(10)}$$

Since all inputs are functions of time, I can decompose the growth of skill premium into several terms. Let g_x denote the growth rate of x:

$$g_{\pi} = (1-\rho)(g_{h_u} - g_{h_s}) + \rho(g_{\psi_s} - g_{\psi_u}). \qquad \textbf{(11)}$$

The component associated with the growth of the quantity of unskilled relative to skilled labor, $(1-\rho)(g_{h_u} - g_{h_s})$, is the relative quantity effect, and the component associated with the growth of the quality of skilled relative to unskilled labor, $\rho(g_{\psi_s} - g_{\psi_u})$, is the relative efficiency effect.

The signaling mechanism considered in this paper provides a structural interpretation of the movement of the relative efficiency effect. The relative efficiency of skilled labor, or the ratio of the expected talent of the skilled to that of the unskilled labor, increases to reflect the compositional change of the two types of labor following an increase in the enrollment rate, as is illustrated in the static model. As long as the elasticity of substitution parameter ρ is positive, so that

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

the skilled and unskilled labor are gross substitutes, the increase in the relative efficiency of the skilled labor will help generate at least part of the skill premium.

Equilibrium.

DEFINITION (Equilibrium). *An equilibrium of this economy is a list of consumption, capital stock, and enrollment status* $\{c(t;\theta,k_0), k(t;\theta,k_0), e(t;\theta,k_0)\}_{t=0}^{\infty}$ *for each cohort-t agent from dynasty* (θ, k_0) *and a list of prices* $\{R(t), \overline{W}(t), \underline{W}(t)\}_{t=0}^{\infty}$, *given the initial distribution of capital* $F(\cdot)$ *over* $[0, \overline{k}_0]$ *and the distribution of talent* $G(\cdot)$ *over* $[0, \overline{\theta}]$, *the positive rate of intergenerational transfer* ϕ *and the production technology* $Y(k,u,s)$ *so that (i) all agents optimally make schooling decision*

$$e(k(t;\theta,k_0)) = \begin{cases} 1, & \text{if } (\theta,k_0) \text{ attends college at } t \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

(ii) firms maximize the current period profit, and (iii) factor markets clear: for all t,

$$K(t) = \int_0^{\overline{\theta}} \int_0^{\overline{k}_0} [k(t;\theta,k_0) - e(k(t;\theta,k_0))Q]\, dF(k_0)\, dG(\theta); \tag{13}$$

$$1 = h_s(t) + h_u(t). \tag{14}$$

I focus on a type of the equilibrium that is separating only in terms of the initial wealth. Call it a *wealth-separating equilibrium*. Along the path of a wealth-separating equilibrium, whose existence I will shortly turn to, as the economy grows and agents accumulate capital, the selection effect from wealth on schooling will bring about changes in the average efficiency units of skilled and unskilled labor, contributing to the dynamics of skill premium. In a wealth-separating equilibrium, all agents optimally go to college as soon as college becomes feasible (that is, as soon as the current capital holdings exceed the cost of college):

$$e(k(t;\theta,k_0)) = \begin{cases} 1, & k(t;\theta,k_0) \geq Q \\ 0, & k(t;\theta,k_0) < Q \end{cases} \tag{15}$$

One immediate implication from Lemma 2 is that, in the wealth-separating equilibrium, since the schooling decision depends only on family wealth, agents from initially wealthier dynasties start going to college earlier than agents from initially poorer dynasties. Moreover, once a dynasty starts attempting college, it will keep on doing so and hence there will be a growing fraction of agents attending college.

Let $x(t)$ denote the fraction of agents going to college at time t. Let $\widehat{k_0}(t)$ denote the initial wealth endowment of the dynasty whose agent attends college for the first time at t. Lemma 2 and (15) imply

$$x(t) = 1 - F(\widehat{k_0}(t)). \tag{16}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

In words, the college enrollment rate at time t is the fraction of dynasties whose initial wealth is above that of a dynasty who sends its agent to college for the first time at t. Given the environment and the definition of the wealth-separating equilibrium, I can write the quantity and efficiency of skilled and unskilled labor for a given enrollment rate $x(t)$ as

$$h_s(t) = x(t)\int_0^{\overline{\theta}} p(\theta)dG; \tag{17}$$

$$\psi_s(t) = E_t(\theta|s) = \frac{\int_0^{\overline{\theta}} \theta p(\theta)dG}{\int_0^{\overline{\theta}} p(\theta)dG}; \tag{18}$$

$$h_u(t) = 1 - x(t)\int_0^{\overline{\theta}} p(\theta)dG; \tag{19}$$

$$\psi_u(t) = E_t(\theta|u) = \frac{\int_0^{\overline{\theta}} \theta dG - x(t)\int_0^{\overline{\theta}} \theta p(\theta)dG}{1 - x(t)\int_0^{\overline{\theta}} p(\theta)dG}. \tag{20}$$

The quantity of skilled labor (h_s) is the proportion of the college attendees who eventually complete college, while the quantity of unskilled labor (h_u) is the remaining population. The efficiency unit of the skilled labor (ψ_s) is the expected talent of college graduates, while the efficiency unit of the unskilled labor (ψ_u) is the expected talent of high school graduates.

It is clear that the supply of skilled labor increases and that of unskilled labor decreases whenever the enrollment rate increases; on the other hand, the efficiency of skilled labor remains constant and that of unskilled labor deteriorates whenever the enrollment rate increases. In other words, the relative quantity effect, $(1 - \rho)(g_{h_u} - g_{h_s})$, is negative and the relative efficiency effect, $\rho(g_{\psi_s} - g_{\psi_u})$, is positive for any elasticity of substitution parameter ρ greater than 0 and on any path with increasing enrollment rates. If I find an environment in which the schooling decision (15) is indeed optimal under the equilibrium factor prices and the relative efficiency effect dominates the relative quantity effect, then I have constructed a wealth-separating equilibrium in which both the skill premium and the college enrollment rate increases over time. This is achieved in Proposition 1.

PROPOSITION 1. *For sufficiently high ρ, and sufficiently low λ and Q, there exists a wealth-separating equilibrium where the college enrollment rate increases together with the skill premium.*

The exact restrictions on the parameters can be found in the proof, but a few comments on the restrictions are warranted here. To ensure that the skilled labor is always paid a higher wage than the unskilled labor, I need the share of output contributed by unskilled labor, λ, to be sufficiently small relative to the share of output contributed by skilled labor, $1 - \lambda$. This translates into the parametric

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

restriction on the share parameter λ:

$$\lambda < \frac{1}{\left(\frac{h_s(0)}{h_u(0)}\right)^{1-\rho}\left(\frac{\psi_u(0)}{\psi_s(0)}\right)^{\rho}+1}. \tag{21}$$

The parametric restriction (21) is obtained from the requirement that the skill premium at the initial instant $\pi(0) > 1$. Since the skill premium increases along the constructed equilibrium path, it suffices to ensure that the initial skill premium is greater than 1.

To ensure that the skill premium increases in the enrollment rate, I need the elasticity of substitution between the skilled and unskilled labor, $\frac{1}{1-\rho}$, to be sufficiently large (and necessarily larger than 1):

$$\frac{1}{1-\rho} \geq 1 + \frac{\int_0^{\overline{\theta}} \theta dG - x(0)\int_0^{\overline{\theta}} \theta p(\theta) dG}{x(0)(\int_0^{\overline{\theta}} \theta p(\theta) dG - \int_0^{\overline{\theta}} p(\theta) dG \int_0^{\overline{\theta}} \theta dG)}, \tag{22}$$

or,

$$\frac{1}{1-\rho} \geq 1 + \frac{h_u(0)}{h_s(0)} \Bigg/ \left[\frac{E(\theta|s)}{E_0(\theta|u)} - \frac{E(\theta)}{E_0(\theta|u)}\right]. \tag{23}$$

The parametric restriction (22) requires the rate of change of the skill premium π in the enrollment rate x to be positive at the initial instant. Along the equilibrium path, the enrollment increases, which relaxes the inequality (22). Rewriting the condition into (23) also suggests that if the college is effective in sorting out high talents, in the sense that the probability of college completion $p(\theta)$ rises steeply at larger θ, then it is more likely that the expected talent of college graduates $E(\theta|s)$ is much bigger than the average talent $E(\theta)$ and the expected talent of high school graduates $E(\theta|u)$ is relatively low. In this case, the relative efficiency effect can overcome the relative quantity effect for relatively low substitution elasticities. In the next section, I examine formally the implication of various values of the parameter ρ on the evolution of the skill premium within the structure of the model.

For any value of the share parameter λ and the elasticity of substitution parameter ρ satisfying the above restrictions, I can find an upper bound of college cost, $\widehat{Q}$, so that, as long as the cost of college in the model falls below $\widehat{Q}$, agents of all talent find attending college optimal. In this wealth-separating equilibrium, the change in the skilled and unskilled labor supply is governed by the change in the enrollment rate, which is in turn pinned down by the cutoff in the initial wealth, $\widehat{k}_0(t)$. Therefore, the equilibrium path can be completely characterized by a dynamic system in two variables, the aggregate capital, $K(t)$, and the cutoff wealth level, $\widehat{k}_0(t)$:

$$\begin{cases} \dot{K}(t) = \phi Y\big(K(t) - x(t)Q, 1 - x(t)\int_0^{\overline{\theta}} p(\theta)dG, x(t)\int_0^{\overline{\theta}} p(\theta)dG\big) \\ \dot{\widehat{k}}_0(t) = -\phi[R(t)Q + \underline{W}(t)] \end{cases} \tag{24}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

$$\text{where } x(t) = 1 - F(\widehat{k_0}(t)),$$
$$\text{s.t. } \widehat{k_0}(t) \geq 0, \text{ with } K(0) = \int_0^{\overline{k}_0} k_0 dF(k_0) \text{ and } k_0(0) = Q.$$

The first equation in the dynamic system says that the increase in aggregate capital comes from the positive transfer of the current generation to the next, which is a fraction ϕ of the current output. The second equation describes how fast the initially poor dynasties start sending their agents to school. More specifically, the threshold in terms of the initial wealth of the marginal dynasty drops exactly by the intergenerational transfer the unskilled is able to make.

2.3. A Theoretical Bound of the Effect of the Education Signal

The idea of the education signal is well known and dates back to the seminal contributions by Spence (1973) and Stiglitz (1975). The dynamic equilibrium formalized in this paper builds on their theoretical insight. It is, however, not easy, in empirical work, to isolate the contribution to the college wage premium from the signaling component and the human capital production component of a college education [Riley (2001), Taber (2001), Fang (2006)]. In my framework, the growth in the skill premium can be decomposed into two components, the relative quantity effect and the relative efficiency effect [recall (11) in Section 2.2]. The relative quantity effect concerns the impact on the skill premium of a decline in the number of skilled workers relative to unskilled workers, while the relative efficiency effect concerns the impact on the skill premium of an increase in the average talent of skilled workers relative to unskilled workers. In this section, I seek to place bounds on how much of the growth in the skill premium can be attributed to the relative efficiency effect that is driven by the sorting patterns in the wealth-separating equilibrium, which I interpret as the signaling channel.

The relative efficiency effect, $g_{\psi_s} - g_{\psi_u}$, or the ratio of the expected talent of a college graduate to the expected talent of a high school graduate, has a closed form in my model. I then choose the underlying parameters, in particular the distribution of talent $G(\cdot)$ and the probability of college completion $p(\cdot)$, to maximize (or minimize) it:

$$\sup_{\substack{G_t(\cdot)\\ p_t(\cdot)}} (or \;\; \inf) \quad g_{\psi_s} - g_{\psi_u}$$

$$\equiv \frac{\int_0^{\overline{\theta}} \theta p_t(\theta) dG_t - \int_0^{\overline{\theta}} p_t(\theta) dG_t \int_0^{\overline{\theta}} \theta dG_t}{(1 - x(t) \int_0^{\overline{\theta}} p_t(\theta) dG_t)(\int_0^{\overline{\theta}} \theta dG_t - x(t) \int_0^{\overline{\theta}} \theta p_t(\theta) dG_t)} \dot{x}(t). \qquad \textbf{(25)}$$

The result is presented in the following proposition.

PROPOSITION 2. *Let the average completion rate* $\int_0^{\overline{\theta}} p_t(\theta) dG_t$ *be bounded from below by* η *and the ratio of the average talent of college graduates and the population average talent be bounded from below by* ξ *(greater than* 1*).*

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

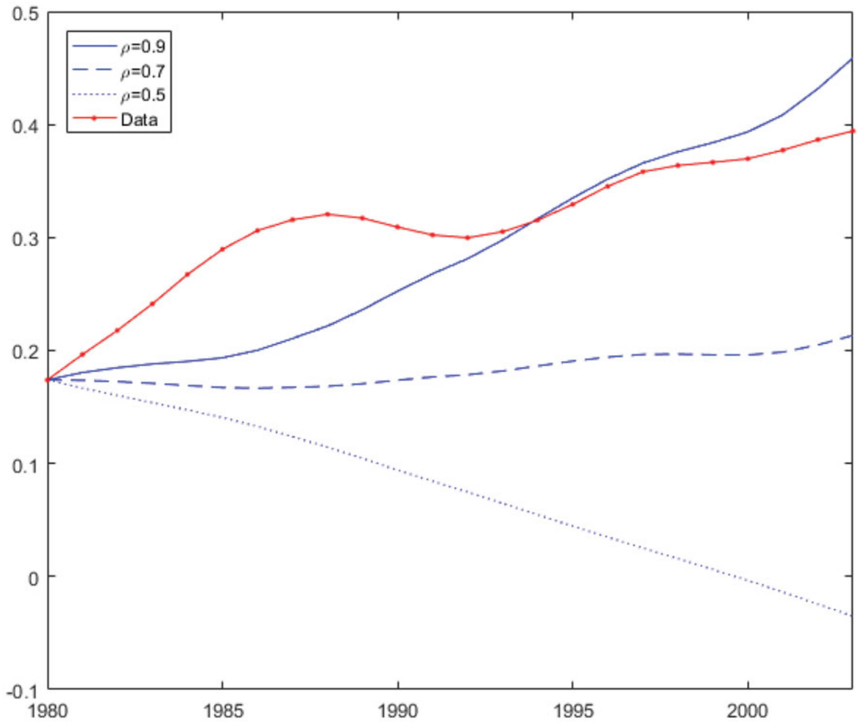

FIGURE 2. Skill premium generated by the model with max relative efficiency effect but without residual SBTC for high values of the elasticity of substitution parameter ρ. The blue lines are the (logged) skill premium generated from a model with maximal relative efficiency effect and no residual trend in SBTC, or the model (27), for different values of ρ.

Suppose $\eta\xi < 1$. The relative efficiency effect, $\rho(g_{\psi_s} - g_{\psi_u})$, in a wealth-separating equilibrium is bounded by the following inequalities:

$$\rho \frac{\eta(\xi - 1)\dot{x}}{(1 - x\eta)(1 - x\eta\xi)} \leq \rho(g_{\psi_s} - g_{\psi_u}) \leq \rho \frac{\dot{x}}{1 - x} = -\rho g_{1-x}. \quad \textbf{(26)}$$

Notice that the upper bound on the relative efficiency effect is expressed in the observables only. This gives us a convenient way to gauge whether the compositional change alone can be the driving force of the growth in skill premium. For that to be true, the relative efficiency effect has to be big enough to overcome the relative quantity effect. The answer to this question is negative for empirically plausible elasticities of substitution between skilled and unskilled labor. In Figure 2, I plot the logged skill premium generated by the model with the maximal relative efficiency effect, $-\rho g_{1-x}$, for the elasticity of substitution parameter ρ

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

equal to 0.5, 0.7, and 0.9:

$$(1-\rho)(g_{h_u}-g_{h_s})-\rho g_{1-x}, \quad \text{for } \rho = 0.5, 0.7, 0.9, \tag{27}$$

where the quantities of skilled and unskilled labor are calculated as in (17) and (19) with the US data of the college enrollment rates and the college completion rates.[8] With a substitution elasticity of 3.3 (or $\rho = 0.7$), the relative efficiency effect, at the maximum, is just strong enough to overcome the relative quantity effect. With a substitution elasticity of 2 (or $\rho = 0.5$), even with the maximal relative efficiency effect, the skill premium predicted from the model is declining.

Since the signaling mechanism by itself cannot generate the entire increase in the skill premium for empirically plausible ρ ranging from 0.3 to 0.5, I allow an additional residual trend in the relative efficiency effect representing the SBTC. I fit the skill premium to data by choosing the growth of the SBTC, g_{SBTC1} (g_{SBTC2}), assuming the maximum (minimum) signaling effect:

$$g_{\pi 1} = (1-\rho)(g_{h_u}-g_{h_s})+\rho(-g_{1-x}+g_{\text{SBTC1}}), \tag{28}$$

$$g_{\pi 2} = (1-\rho)(g_{h_u}-g_{h_s})+\rho\left[\frac{\eta(\xi-1)\dot{x}}{(1-x\eta)(1-x\eta\xi)}+g_{\text{SBTC2}}\right]. \tag{29}$$

To compute the minimum signaling effect, I set the lower bound of the college completion rate η to be the lowest college completion rate over the sample period in the data, which is 0.63, and set the lower bound of the ratio of the average talent of college graduates to the population average ξ to be the lowest ratio between the wage of skilled labor and the average wage over the sample period in the data, which is 1.14.

Once I have the residual SBTC g_{SBTC1} and g_{SBTC2}, I simulate a series of skill premium, shutting down the signaling channel entirely:

$$\widehat{g}_{\pi 1} = (1-\rho)(g_{h_u}-g_{h_s})+\rho g_{\text{SBTC1}}, \tag{30}$$

$$\widehat{g}_{\pi 2} = (1-\rho)(g_{h_u}-g_{h_s})+\rho g_{\text{SBTC2}}, \tag{31}$$

Let the skill premium grow from the initial level in the data at rates $\widehat{g}_{\pi 1}$ and $\widehat{g}_{\pi 2}$, which produces two counterfactual skill premium series, $\{\widehat{\pi}_{1,t}\}$ and $\{\widehat{\pi}_{2,t}\}$.

The upper (lower) bound of the signaling effect is then measured by the distance between the growth in the data skill premium, $\{\text{skpm}_t\}$, and the growth in the counterfactual skill premium, $\{\widehat{\pi}_{1,t}\}$ ($\{\widehat{\pi}_{2,t}\}$):

$$\left(1-\frac{\log(\widehat{\pi}_{iT})-\log(\widehat{\pi}_{i1})}{\log(\text{skpm}_T)-\log(\text{skpm}_1)}\right)\times 100, \text{ for } i = 1, 2\ldots, \tag{32}$$

Figure 3 shows the maximum and minimum signaling effects for the elasticity of substitution parameter ρ ranging from 0.3 to 0.5. The effect of signaling can range from just over 10% of the observed growth in skill premium to over 60%

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

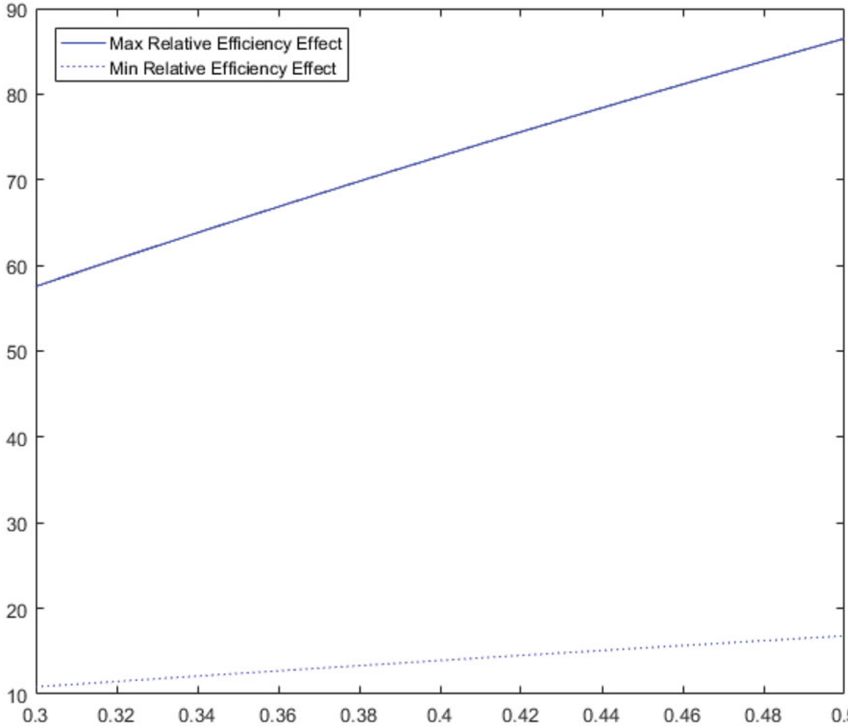

FIGURE 3. Percentage of skill premium driven by changing relative efficiency effect for various elasticity of substitution parameter ρ. This is the percentage of the (logged) skill premium that is generated by a changing relative efficiency effect, a measure of the signaling effect of this model, for different values of ρ.

for a ρ of 0.4, the value chosen for the calibration exercise in Section 3. The exact magnitude of the signaling effect is the subject of Section 3.

2.4. Alternative Sorting Patterns

There are two main theoretical deviations from the wealth-separating equilibrium that I focus on in this paper. I discuss them here.

The first deviation is from the assumption that the initial wealth and talent are uncorrelated, while still maintaining an equilibrium separating only in terms of wealth. The interesting case to study is when wealth and talent are positively correlated. In this case, a relaxation of the budget constraint can lead to different compositional changes in the labor force and movements in the skill premium.

Consider the static model in Section 2.1. Assume instead that the initial wealth and talent are perfectly positively correlated. Let the median initial wealth level be

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

denoted as k_m. Then, all those whose wealth is below (above) k_m have low (high) talent:

$$\theta = \begin{cases} \underline{\theta}, & \text{if } k \leq k_m \\ \overline{\theta}, & \text{if } k > k_m. \end{cases}$$

Suppose $Q \geq k_m$, so only the few highly talented go to college, and then the skilled wage $\overline{W} = \overline{\theta}$. Let $x \equiv 1 - F(Q) \leq \frac{1}{2}$ be the fraction of agents that go to college. The pool of high school graduates then consists of a measure 0.5 of low talent who cannot afford college, a measure $0.5 - x$ of high talent who cannot afford college, and a measure $(1 - \overline{p})x$ of high talent who fails college. As x increases, one observes a net flow of the high talent from the unskilled to the skilled labor force, which drives the unskilled wage down. The wage gap therefore increases.

Suppose $Q < k_m$ and the majority attends college. The pool of college graduates then consists of a measure $0.5\overline{p}$ of high talent and a measure $(x - 0.5)\underline{p}$ of low talent agents. As x increases, it draws in low talent agents in the skilled labor force, lowering the average wage for skilled labor. The pool of high school graduates consists of a measure $(1 - x)$ of low talent who cannot afford college, a measure $(x - 0.5)(1 - \underline{p})$ of low talent who fails college and a measure $0.5(1 - \overline{p})$ of high talent who fails college. Now, an increase in x pushes the low talent to attempt college, some of which will actually get the college degree, resulting in a net flow of low talent from the unskilled to the skilled labor force. This drives up the unskilled wage. The wage gap consequently decreases.[9]

Under the assumption of a perfectly positive correlation between initial wealth and talent, the skill premium first increases and then decreases, as college enrollment grows from 0 to 1. When there are relatively few attending college, the way the observables behave in this setting is very similar to that under zero correlation between initial wealth and talent. In either case, the average talent of the skilled is flat and the average talent of the unskilled declines, as enrollment increases. Assuming a positive correlation tends to increase the rate at which the average talent of the unskilled declines in college attendance, since it is the high talent that is being moved at the margin.[10] By adopting the assumption of a zero correlation, I am providing a conservative estimate of the signaling force. The theoretical possibility that when college becomes such a mass phenomenon that increased enrollment closes the wage gap, however, seems empirically implausible. Hendricks and Schoellman (2014) show that the ability gap between high school graduates and college graduates increased between the 1910 and 1960 cohorts, and the increasing gap is driven mostly by a decrease in the average ability of the high school only.

The second deviation from the wealth-separating equilibrium is an equilibrium that is separating in both wealth and talent. In Appendix, I prove the existence of such a dynamic equilibrium for a special case of the CES production where skilled and unskilled labor are perfect substitutes. In this equilibrium, all agents

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

whose initial dynastic capital is above some threshold $\widehat{k_0}(t)$ and whose talent is above some threshold $\widehat{\theta}(t)$ go to college. Along the equilibrium path, the skill premium, however, is not necessarily monotone. A sufficient condition to ensure the simultaneous increase in college enrollment and skill premium is to have a sufficiently high threshold for talent. In that world, both the average talent of college graduates and high school graduates declines, but that of high school graduates declines even more. I choose to focus on the wealth-separating equilibrium because it offers sharper and cleaner theoretical predictions and requires fewer assumptions on the unobservable talent. More pragmatically, the starting point of the model is a group of high school seniors—since I do not consider high school dropouts, one would think that high school students who work through their senior year must have believed that they have a chance at college and college is beneficial to them [Bedard (2001)].

3. A QUANTITATIVE ASSESSMENT: US 1980–2003

During the period 1974 to 1997, the United States saw a monotonically increasing trend in the college enrollment rates, which provides us with a convenient environment to evaluate the signaling mechanism developed in this paper (Figure 1). Section 2.3 makes it clear that the signaling mechanism is unlikely to be the sole driver of the increase in skill premium. Therefore, the proper question to ask is, under reasonable parameter values, how much the signaling mechanism can help generate the observed skill premium. To that end, I estimate a semi-reduced-form model in the spirit of (28) or (29), but retaining the structural elements in the relative efficiency effect.

To be more specific, I incorporate the reduced-form SBTC, a demand-side explanation, into the model by replacing the evolution of the efficiency unit of skilled labor $\psi_s(t)$ in (18) with[11]

$$\psi_s(t) = e^{(1+\gamma_{\text{SBTC}})\cdot t}\,\frac{\int_0^{\overline{\theta}} \theta p(\theta) dG}{\int_0^{\overline{\theta}} p(\theta) dG}. \tag{33}$$

The parameter γ_{SBTC} in $\psi_s(t)$ is the residual growth rate of SBTC that is needed on the top of the contribution from the signal to generate the skill premium.

In order to assess the contribution of the signaling mechanism, I first fit the model to the data (according to a procedure specified later). Then, I fix the enrollment rate in the efficiency unit of the unskilled labor $\psi_u(t)$ at its initial level, x_0:

$$\psi_u(t) = \frac{\int_0^{\overline{\theta}} \theta dG - x_0 \int_0^{\overline{\theta}} \theta p(\theta) dG}{1 - x_0 \int_0^{\overline{\theta}} p(\theta) dG}, \tag{34}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

and simulate a hypothetical trend of skill premium. Under a stationary environment where the distribution of talent $G(\cdot)$ and the college completion probabilities $p(\cdot)$ are time invariant, this would imply a fixed efficiency unit of the unskilled labor. The simulated skill premium would, in general, rise more slowly than the predicted skill premium from the fitted model. I interpret the difference in the trend growth of the two series of skill premium as a measure of the signaling effect in the same way as in (32).

The enrollment rates $x(t)$ and the skill premium $\pi(t)$ in the model have straightforward data counterparts. The term $\int_0^{\overline{\theta}} p(\theta)dG$ corresponds to the college completion rate in the population. With the enrollment rate and college completion rate in hand, I can construct the supply of skilled young workers, h_s, and the unskilled young workers, h_u, according to (17) and (19). I develop two alternative strategies to pin down the efficiency units of skilled and unskilled labor and discuss the implications of these strategies on the measure of the signaling effect. In what follows, I discuss the data, the methodology and present the results.

3.1. Data

The data are structured to facilitate the interpretation of a period in the model. The model year refers to the year for which the skill premium is calculated. Within the same period in the model, the enrollment rate of 6 years and the college completion rate of 2 years before the model year are used. This is to accommodate the fact that the skill premium is calculated for the age group of 23–26 years. The first period in our sample is 1980 and the last is 2003.

Skill premium. To be consistent with the theoretical prediction that later cohorts who are subject to a stronger signaling effect face a higher premium, the calculation of college premium is cohort based. I computed the wage series using the current population survey (CPS) March data from 1980 to 2003 by age groups and focus on the age group of 23–26 year olds. Only full-year, full-time workers that have positive wage and schooling are considered. The skill premium is the ratio between the weekly wage of a college graduate and the weekly wage of a high school graduate. In order to compute the wage rates, I regress the reported weekly wage by gender on dummies of education, geographic region, and race. CPS sampling weights are used. The education has five categories: high school dropouts, high school graduates, some college, college graduates, and above. The geographical region has four: Northeast region, Midwest region, South region, and West region. The race variable has three: white, black, and other. The weekly wage of a college graduate (or a high school graduate) is the sample average of the predicted wage for a white worker with a college degree (or with either some college or a high school diploma) across geographical regions. I compute the skill premium by gender and by year. The skill premium in a given year is then obtained by averaging the gender-specific skill premia in that year with the gender-specific aggregate weeks worked as weights. Last, I apply the Hodrick–Prescott (HP) filter

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

with a smoothing parameter of 6.25 to this annual series of the skill premium. The construction of the college wage premium essentially follows Autor et al. (2008).

College enrollment rate. College enrollment and the number of high school completers from 1974 to 1997 are taken from Table 191 "College enrollment and enrollment rates of recent high school completers, by sex: 1960 through 2006" in *Digest of Education Statistics 2007*, available on the National Center for Education Statistics website. The definition of high school completers is all individuals aged 16 to 24 years who graduated from high school or completed a GED during the preceding 12 months. The enrollment rate is the ratio between the total enrollment in a given year over the total number of high school completers. The HP filter with a smoothing parameter of 6.25 is also applied to this series.

College completion rate. To construct the college completion rate, I take the number of bachelor's degrees conferred by degree-granting institutions each year from 1978 to 2001 and divide it by the total college enrollment 4 years before. The number of bachelor's degrees conferred by degree-granting institutions by control of institution and by year is taken from Table 266 "Degrees conferred by degree-granting institutions, by control of institution and level of degree: 1969–70 through 2005–06" in *Digest of Education Statistics 2007*. Only bachelor's degrees are counted; the total number is the sum of the number of degrees conferred by public institutions and by private institutions. The raw data show a clear upward trend as well as a high volatility. I then regress the raw data on the year and the year squared and use the predicted completion rate in the quantitative analysis.

Cost of college. In order to construct the real net cost of college, I need both the sticker price – tuition, fees, room and board (TFRB) – and the total aid per student. The TFRB is reported by types of institution in Table 5 "Average published tuition and fee and room and board charges at 4-year institutions in constant 2009 dollars, 1979–80 to 2009–10" in *Trends in College Pricing 2009*. I take the enrollment-weighted average of the TFRB in public institutions and in private institutions. The average aid per student is taken from the source data of Figure 11 "Average aid per full-time equivalent student in constant (2008) dollars, 1973–74 to 2008–09" in *Trends in Student Aid 2009*. The total aid includes grants, loans (excluding private nonfederal loans), federal work study, and education tax benefits. After converting both the TFRB and the aid in constant 2000 dollars, I define the difference between the two as the real net cost of college. The net cost of college rose sharply from 1980 to 1982 and continued to rise modestly for the rest of the sample years (Figure 4). I fitted a linear time trend in the net price of college, which grows by 24.3% from \$3952 in 1980 to \$4911 in 2003.

Initial income distribution in 1980. The income distribution in the initial period of the model is proxied by the income distribution in 1980. From the CEPR Uniform Extracts of the March CPS 1980, I keep the total family income of married

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

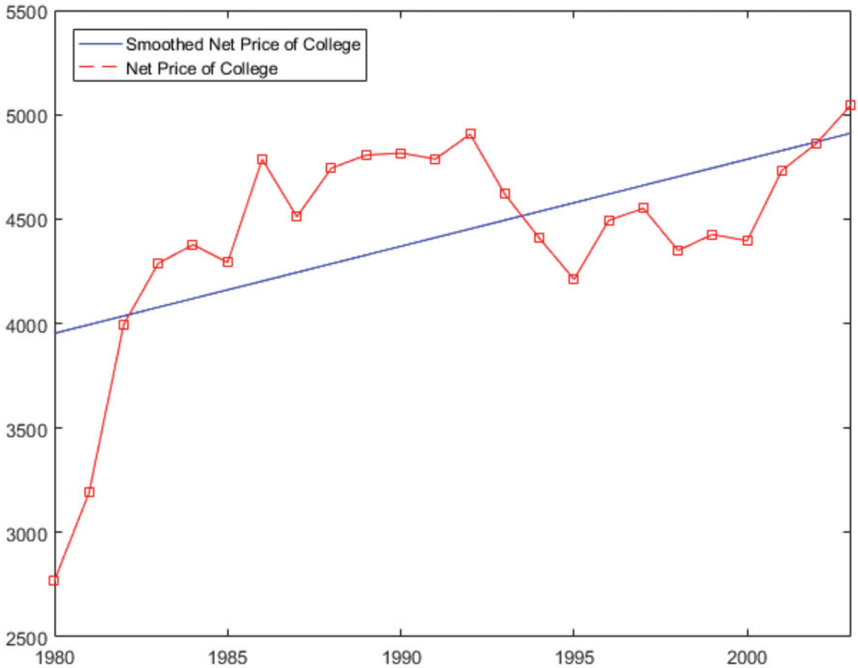

FIGURE 4. Net price of college (in 2000$): Tuition, fees, room and board net total aid per full-time-equivalent (FTE) student. The TFRB is from *Trend in College Pricing 2009* and the total aid per FTE student is from *Trend in Student Aid 2009*. Both are published by the College Board. For the details of the data, see Section 3.1.

males aged between 40 and 50 years, who are likely to have children around 20-year-old facing the decision of whether or not to attend college. I trimmed the top and bottom 1% of the family income to avoid outliers, deflated it to constant 2000 dollars and estimated a kernel density based on a normal kernel function.

In what follows, denote the empirical series of the skill premium, the enrollment rates, and the college completion rates as skpm_t, enrl_t, and comp_t, respectively.

3.2. Methodology

The general procedure consists of two steps. First, I conduct a simple calibration of the model described in (24). To accommodate the time trend in the cost of college, the equation that describes the evolution of the cutoff initial capital $\widehat{k}_0(t)$ in system (24) is modified as follows:

$$\dot{\widehat{k}_0}(t) = -\phi[R(t)Q(t) + \underline{W}(t)] + \dot{Q}(t). \quad \textbf{(35)}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

TABLE 1. Model parameters

Model	Value	Interpretation
ρ	0.4	It implies an elasticity of substitution between skilled and unskilled labor of 1.67 (KORV).
μ	1/3	Income share parameter of capital versus aggregate labor.
σ	-1	It implies an elasticity of substitution between capital and aggregated labor of 0.5. [Antras (2004)]
λ	0.6	Income share parameter of unskilled labor versus skilled labor.

I will verify in the calibrated model with time-varying cost of college that the assumptions of the existence of a wealth-separating equilibrium are satisfied. The calibration proceeds in two nested optimization routines. In the outside loop, I choose the rate of growth of the SBTC, γ_{SBTC}, to minimize the distance between the model generated skill premium and that in the data. In the inside loop, for a given value of γ_{SBTC}, I choose the rate of intergenerational transfer ϕ to minimize the distance between the model generated enrollment rate and that in the data. With the fitted model in hand, in the second step, I simulate the skill premium from this model, keeping the enrollment rates in the efficiency unit of the unskilled labor fixed at the initial enrollment rate, i.e., according to (34). Unsurprisingly, the resulting skill premium grows a lot slower than the skill premium generated by the fitted model. The difference in the growth of these two series of skill premium is attributed to the signaling mechanism.

The key parameters of the model are taken from the estimates from the empirical literature (Table 1).

To construct the efficiency units of skilled and unskilled labor, I need to pin down $\int_0^{\overline{\theta}} \theta dG$ and $\int_0^{\overline{\theta}} \theta p(\theta) dG$. The former measures the aggregate talent in this economy and the latter measures the aggregate talent of skilled labor if all agents attend college, or the maximum aggregate talent of college graduates. Since I effectively allow the college completion rate $\int_0^{\overline{\theta}} p(\theta) dG$ to vary over time by equating it with its empirical counterpart, comp_t, I should also allow the maximum aggregate talent of college graduates $\int_0^{\overline{\theta}} \theta p(\theta) dG$ to vary over time. To the extent that there is no consensus on an adequate measure of talent for the US during the sample period, to which I can directly calibrate, I determine these series within the model.

Ultimately, the value of the aggregate talent $\int_0^{\overline{\theta}} \theta dG$ and the maximum aggregate talent of skilled labor $\int_0^{\overline{\theta}} \theta p_t(\theta) dG$ determine: (1) the relative productivity of skilled versus unskilled labor, and (2) the relative productivity of labor versus capital. Point (1) suggests that I determine the ratio $\int_0^{\overline{\theta}} \theta dG / \int_0^{\overline{\theta}} \theta p_0(\theta) dG$ in the initial period by matching the initial level of skill premium in the model with that

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

in the data:

$$\text{skpm}_0 = \frac{1-\lambda}{\lambda}\frac{h_{u0}}{h_{s0}}\left(\frac{s_0}{u_0}\right)^{\rho}$$
$$= \frac{1-\lambda}{\lambda}\frac{1-\text{enrl}_0\cdot\text{comp}_0}{\text{enrl}_0\cdot\text{comp}_0}\left(\frac{\text{enrl}_0}{\int_0^{\overline{\theta}}\theta dG/\int_0^{\overline{\theta}}\theta p_0(\theta)dG-\text{enrl}_0}\right)^{\rho}. \quad \textbf{(36)}$$

Next, given the ratio $\int_0^{\overline{\theta}}\theta dG/\int_0^{\overline{\theta}}\theta p_0(\theta)dG$, Point (2) suggests that I solve the maximum aggregate talent of college graduates in the initial period $\int_0^{\overline{\theta}}\theta p_0(\theta)dG$ by taking the capital–labor share of income in the initial period to be $1/2$:

$$\frac{1}{2} = \frac{\mu}{1-\mu}\left\{\frac{k_0}{[\lambda u_0^{\rho}+(1-\lambda)s_0^{\rho}]^{\frac{1}{\rho}}}\right\}^{\sigma} \quad \textbf{(37)}$$
$$= \frac{\mu}{1-\mu}\left\{\frac{k_0}{\int_0^{\overline{\theta}}\theta p_0(\theta)dG[\lambda(\int_0^{\overline{\theta}}\theta dG/\int_0^{\overline{\theta}}\theta p_0(\theta)dG-x_0)^{\rho}+(1-\lambda)x_0^{\rho}]^{\frac{1}{\rho}}}\right\}^{\sigma}$$
$$= \frac{\mu}{1-\mu}$$
$$\times\left\{\frac{\text{mean}(cdf(\cdot))-\text{enrl}_0\cdot Q}{\int_0^{\overline{\theta}}\theta p_0(\theta)dG[\lambda(\int_0^{\overline{\theta}}\theta dG/\int_0^{\overline{\theta}}\theta p_0(\theta)dG-\text{enrl}_0)^{\rho}+(1-\lambda)\cdot\text{enrl}_0^{\rho}]^{\frac{1}{\rho}}}\right\}^{\sigma}.$$

Once I have $\int_0^{\overline{\theta}}\theta p_0(\theta)dG$, it is straightforward to use the ratio to back out the aggregate talent $\int_0^{\overline{\theta}}\theta dG$.

Starting from the maximum aggregate talent of college graduates in the initial period $\int_0^{\overline{\theta}}\theta p_0(\theta)dG$, I consider two alternative models of how this maximum aggregate talent evolves over time.

Model 1 (constant expected talent of college graduates). I let the maximum aggregate talent of college graduates grow at the same rate as the college completion rates:

$$\int_0^{\overline{\theta}}\theta p_t(\theta)dG = \int_0^{\overline{\theta}}\theta p_0(\theta)dG\cdot\frac{\text{comp}_t}{\text{comp}_0}. \quad \textbf{(38)}$$

Model 1 respects the tight restriction from the theory that the expected talent of a college graduate remains constant over time:

$$E[\theta|CG] = \frac{\int_0^{\overline{\theta}}\theta p_t(\theta)dG}{\text{comp}_t} = \frac{\int_0^{\overline{\theta}}\theta p_0(\theta)dG}{\text{comp}_0}. \quad \textbf{(39)}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

Let the skill premium generated by the fitted model be denoted π_t. With the fitted model, I simulate the hypothetical trend of the skill premium, fixing the enrollment rates at the initial level:

$$\widehat{\pi}_t = \frac{1-\lambda}{\lambda}\left(\frac{h_{ut}}{h_{st}}\right)^{1-\rho}\left(\frac{\psi_{st}}{\widehat{\psi}_{ut}}\right)^{\rho}, \tag{40}$$

where the quantity of unskilled and skilled labor, h_{ut} and h_{st}, as well as the efficiency unit of skilled labor ψ_{st} remain as before and the efficiency unit of unskilled labor fixing the enrollment rate at its initial level $\widehat{\psi}_{ut}$ is given by

$$\widehat{\psi}_{ut} = \frac{\int_0^{\overline{\theta}} \theta dG - x_0 \int_0^{\overline{\theta}} \theta p_t(\theta) dG}{1 - x_0 \cdot \mathrm{comp}_t}. \tag{41}$$

The difference between $\widehat{\pi}_t$ and π_t defines the measure of the signaling component.

In light of the recent finding by Carneiro and Lee (2011), requiring the quality of college graduates to remain constant, because of a parsimonious theoretical model, can be somewhat restrictive. Carneiro and Lee show that the average quality of US college graduates has decreased over the 1960 to 2000 period and the decline has a substantial impact on the evolution of the wage distribution. The decline of the quality of college graduates could be due to a competing compositional effect or to lower human capital production in the college sector caused by, for example, congestion. The fact that they find that college enrollment predicts a lower wage to college graduates but has an insignificant effect on the wage to high school graduates seems to support the latter interpretation. In the next model, I take the stand that the declining quality reflects purely a decline in the human capital productivity in the college sector, a supply-side explanation of the evolution of the skill premium in addition to the signaling mechanism I model. I allow for a separate trend for the expected talent of college graduates and calibrate it to match Carneiro and Lee's result that, if the quality of college graduates had been fixed at the 1960 level, the (logged) skill premium would have grown 30% more by 2000.

Model 2 (declining expected talent of college graduates). I calibrate the decline in the expected talent of college graduates so that with the fitted model, if I had fixed the expected talent of the college graduate at the initial level, the model would generate a trend of the (logged) skill premium that is 17.25% higher than the fitted model generates.[12] More precisely, I let the maximum aggregate talent of college graduates vary over time at the rate ω:

$$\int_0^{\overline{\theta}} \theta p_t(\theta) dG = \int_0^{\overline{\theta}} \theta p_0(\theta) dG \cdot (1+\omega)^t, \tag{42}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

where the growth rate ω is lower than the implied growth rate in the series comp_t so that the average quality of a college graduate decreases over time:

$$E_t[\theta|CG] = \int_0^{\overline{\theta}} \theta p_0(\theta) dG \cdot \frac{(1+\omega)^t}{\text{comp}_t}. \quad (43)$$

For each trial of ω, I fit the model as before by first choosing the intergenerational transfer rate ϕ to fit the enrollment data (for a given growth rate of SBTC) and next choosing the growth of SBTC γ_{SBTC} to fit the skill premium data. With this model, I simulate the skill premium, fixing the quality of college graduates $E_t[\theta|CG]$ at the initial level, $\int_0^{\overline{\theta}} \theta p_0(\theta) dG/\text{comp}_0$:

$$\widetilde{\pi}_t = \frac{1-\lambda}{\lambda} \left(\frac{h_{ut}}{h_{st}}\right)^{1-\rho} \left(\frac{\widetilde{\psi}_{st}}{\widetilde{\psi}_{ut}}\right)^{\rho}; \quad (44)$$

where the quantities of the two types of labor h_{ut} and h_{st} remain as before but the efficiency units $\widetilde{\psi}_{st}$ and $\widetilde{\psi}_{ut}$ are

$$\widetilde{\psi}_{st} = (1+\gamma_{\text{SBTC}})^t \frac{\int_0^{\overline{\theta}} \theta p_0(\theta) dG}{\text{comp}_0} \quad (45)$$

$$\widetilde{\psi}_{ut} = \frac{\int_0^{\overline{\theta}} \theta dG - x_t \frac{\int_0^{\overline{\theta}} \theta p_0(\theta) dG}{\text{comp0}} \text{comp}_t}{1 - x_t \cdot \text{comp}_t} \quad (46)$$

The growth parameter in the expected talent of college graduates ω is chosen such that the (logged) skill premium without the decline in the quality of college $\log(\widetilde{\pi}_t)$ is 17.25% higher than the logged skill premium generated by the fitted model. With the choice of ω, I have pinned down all the parameters in the model. Then, I follow the same procedure as before to measure the contribution from the signaling.

Compared with the previous model with a constant expected talent of college graduates (Model 1), I expect it to be more difficult for this model to generate the skill premium, due to the extra downward pressure on the skill premium from the deteriorating quality of the college graduates. Indeed, the implied residual growth of SBTC is higher for this model than for the previous model.

In both models, I also verify that the calibrated models are consistent with the assumptions for the existence of a wealth-separating equilibrium. In particular, the wealth-separating equilibrium requires all talent to (weakly) prefer attending college, regardless of family background. In other words,

$$p(0)(\overline{W}_t - \underline{W}_t) \geq R_t Q_t \quad (47)$$

has to hold for all periods in the calibrated models. I solve out the minimum value of $p(0)$ that satisfies (47) period by period from the calibrated models, and this value ranges from 8% to 11%. This means, a high school senior with the lowest

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

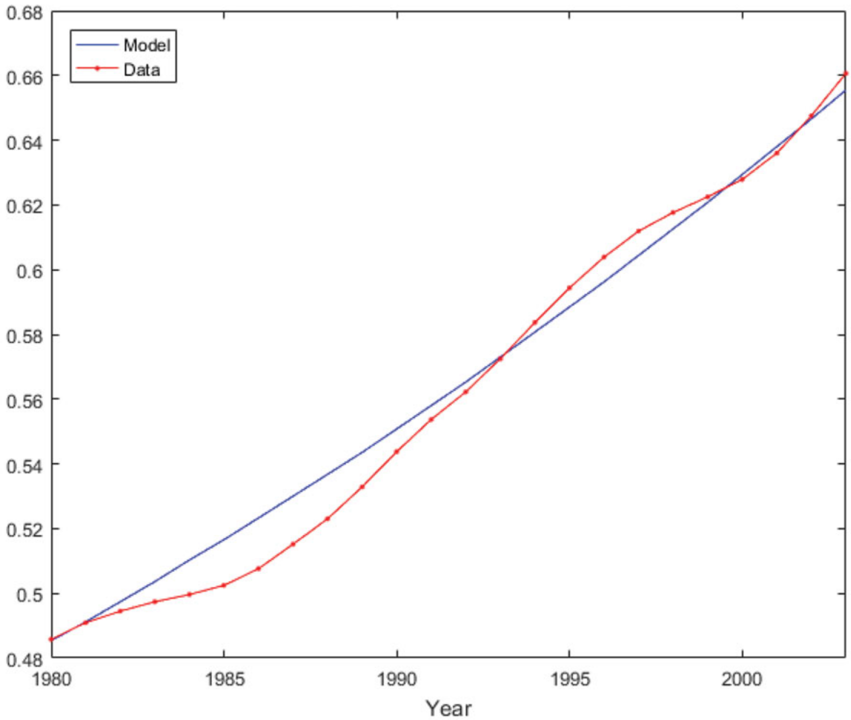

FIGURE 5. Model 1 (constant expected talent of college graduates): Enrollment rates, model vs. data. The solid blue line is the endogenous enrollment rates generated by the fitted Model 1.

talent still has about 10% chance of getting into some college in the calibrated models. This seems to be plausible, given that in reality high school students also have the option of dropping out.

3.3. Results

The fit of the model with a constant expected talent of college graduates (Model 1), in terms of the enrollment rates and the (logged) skill premium, is reported in Figures 5 and 6. In Figure 6, the solid blue line is the skill premium generated by the fitted model over the sample period. Since the model is essentially a theory about the trend of the skill premium, it does not catch the swing in the skill premium in the late 1980s, yet it is able to match well with the overall increase in the skill premium in the data. The dashed blue line is the prediction of the skill premium from the model, fixing the enrollment rates at the initial level in forming the belief of the expected talent of high school graduates. Without recognizing the signaling that comes from the increased college attendance, the model generates an increase of the (logged) skill premium that is 87.32% of the observed increase,

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

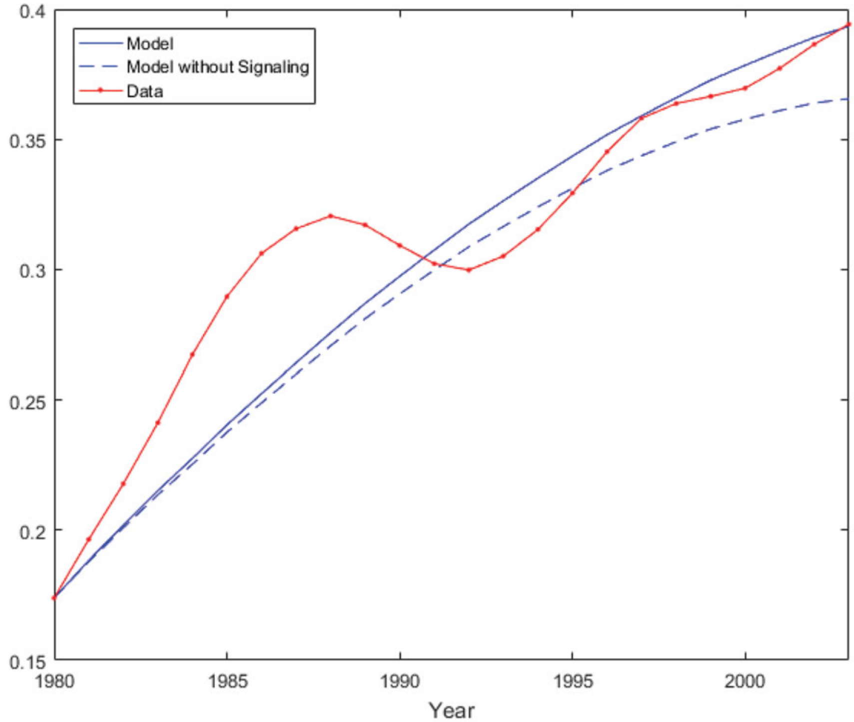

FIGURE 6. Model 1 (constant expected talent of college graduates): Model skill premium with and without signaling. The solid blue line is the (logged) skill premium generated from the fitted Model 1. The dashed blue line is the (logged) skill premium generated from the fitted Model 1 fixing the enrollment rates in the belief of the expected talent of high school graduates at the 1980 level.

or falls 12.68% short of the observed increase. Comparing it to the bounds of the signaling effect in Figure 3, the effect of signaling falls close to the theoretical lower bound I established in Section 2.3.

The fit of the model that allows for a declining expected talent of college graduates (Model 2) is likewise presented in Figures 7 and 8. The implied decline in the quality of college graduates (or the expected talent of college graduates) is 3.85% of the initial level by the end of the 24 years of the sample period. In Figure 9, the dashed blue line shows the simulated skill premium, fixing the quality of the college graduates at the 1980 level. It is 17.25% higher in 2003 relative to the skill premium from the fitted model (the solid blue line). The hypothetical skill premium shutting down the signaling mechanism stands at 91% in 2003 of the skill premium from the fitted model (Figure 8). As expected, the implied growth of SBTC (7.94%) is now higher than in the previous model with a constant expected talent of college graduates (7.66%).

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

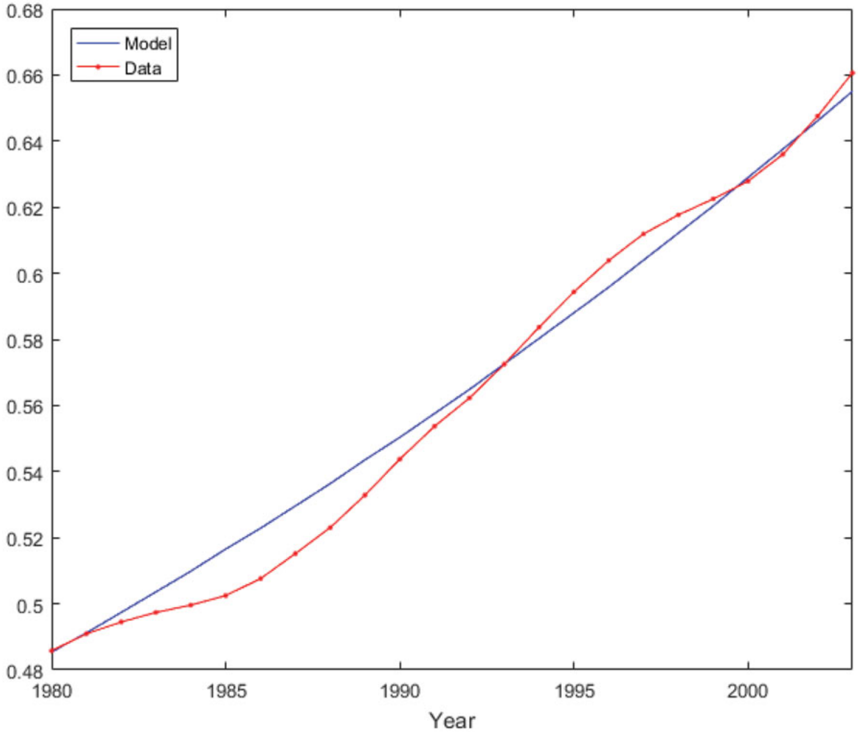

FIGURE 7. Model 2 (declining expected talent of college graduates): Enrollment rates, model vs. data. The solid blue line is the endogenous enrollment rates generated by the fitted Model 2.

I conclude that the effect of the signaling mechanism on the growth of the skill premium is modest. This effect accounts for somewhere from 9.01% to 12.68% of the increase in the observed (logged) skill premium, depending on whether one allows for a declining quality of college education or not.

4. CONCLUSION

Though the idea of education as a job market signal is well known [Spence (1973), Stiglitz (1975)], its application to the evolution of wage distribution has not been well articulated in theory. This paper is such an attempt. I develop a model with dynasties, heterogeneous in initial wealth and talent, who make schooling decisions for successive cohorts of offspring. The growth in the college enrollment rate due to an increased access to college makes a high school diploma a clearer signal of low talent. If talent is useful in production, a college degree will be rewarded a higher premium relative to a high school diploma. This brings about a growing wage gap between college and high school graduates.

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

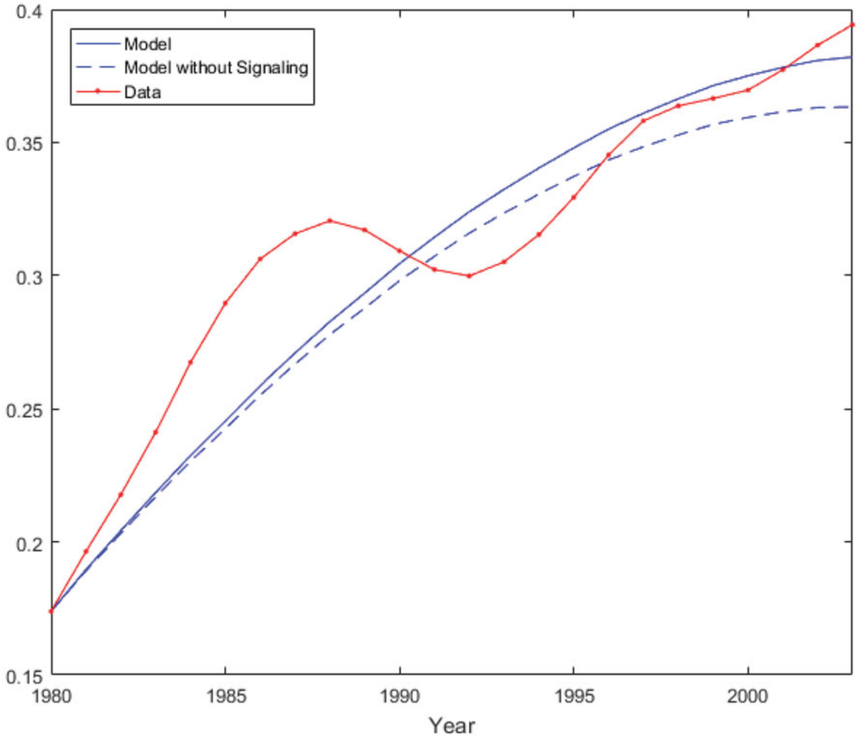

FIGURE 8. Model 2 (declining expected talent of college graduates): Model skill premium with and without signaling. The solid blue line is the (logged) skill premium generated from the fitted Model 2. The dashed blue line is the (logged) skill premium generated from the fitted Model 2 fixing the enrollment rates in the belief of the expected talent of high school graduates at the 1980 level.

I show that the effect from the signaling mechanism that I model tends to be strong when the elasticity of substitution between the college educated and noncollege-educated labor is high and/or when the college sector becomes increasingly efficient in sorting out high talent. However, the theoretical bound on the signaling effect suggests that the signaling mechanism itself is not likely to be the main driving force of the increase of skill premium.

When I calibrate the model to the observed trends in the skill premium for young workers and the college enrollment rates in the United States from 1980 to 2003, I find that the signaling mechanism has a modest but sizeable effect on the increase in the college wage premium. In a model that also adjusts for the declining quality of college graduates, the signaling mechanism accounts for 9.01% of the increase in the observed college wage premium over the 24 years in the sample. I interpret my finding as suggesting that we economists should perhaps take a serious approach to modeling the college sector as the supplier of skilled labor.

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

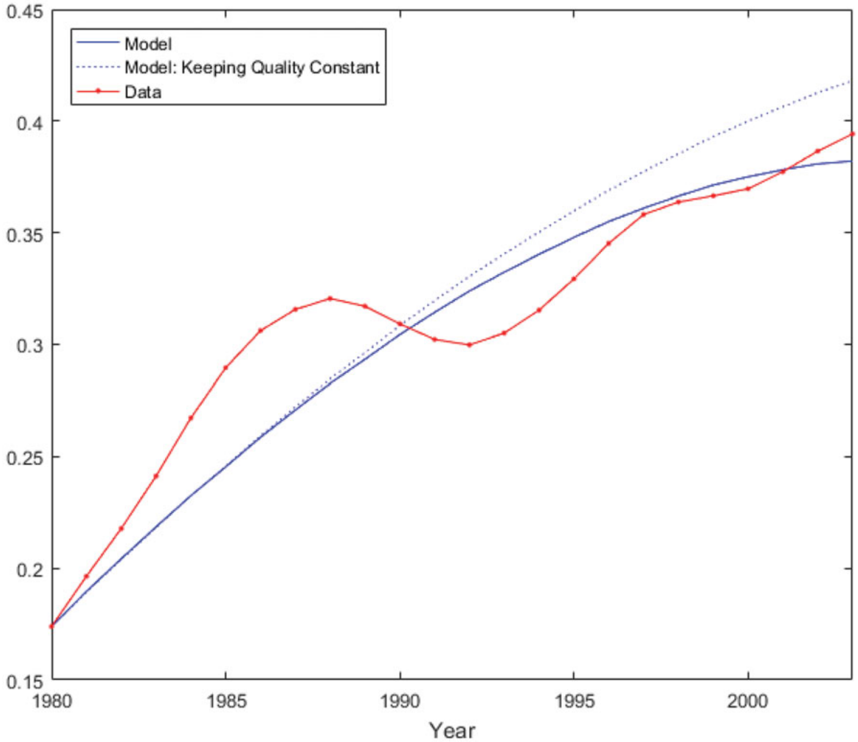

FIGURE 9. Model 2 (declining expected talent of college graduates): Skill premium, skilled labor quality adjusted or not. The solid blue line is the (logged) skill premium generated from the fitted Model 2. The dotted blue line is the (logged) skill premium generated from the fitted Model 2 fixing the expected talent of college graduates in the efficiency units of the skilled and unskilled labor at the 1980 level.

Compositional changes in the "output" of the college sector can imply revisions to what we have understood as the sources of a growing inequality.

NOTES

1. The earnings of unskilled American men saw a dramatic drop in real terms in the 1970s and 1980s [Blackburn et al. (1990)], and did not pick up until mid-1990. This is a period in which the increase in skill premium is driven mostly by a deteriorating wage offer to unskilled labor than an increasing wage offer to skilled labor.

2. Similar intuition is also exploited by Hendel et al. (2005) and Balart (2016), who consider the implication of increased access to the student loan market on the steady-state level of inequality in a dynamic asymmetric information model. In contrast, I consider the price dynamics of skilled and unskilled labor along the entire growth path of a dynamic production model to examine the evolution of inequality.

3. For simplicity, I take talent θ as a time-invariant attribute of a dynasty. The limitation of this simplification is that the model implies zero intergenerational mobility, in the sense that the relative ranking of the wealth of the dynasties is fixed over time in expectation.

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

4. Even though the initial endowments of talent and capital stocks are independent, along the equilibrium path, talent and capital holdings will, in general, be correlated. I discuss the relaxation of this assumption in Section 2.4.

5. The assumption of risk neutrality simplifies the math considerably. However, the main result of the paper does not hinge on this assumption. In a note that is not published and that is available upon request, I show that under mild risk aversion (i.e., for a CRRA parameter in between 0 and 1), the main theoretical result, Proposition 1, continues to hold, provided a more stringent bound on the cost of college. Moreover, the dynamic system that describes the equilibrium paths of enrollment and skill premium continues to be valid.

6. The constant rate of intergenerational transfer ϕ is a common feature in models of intergenerational altruism and human capital investment [e.g., Glomm and Ravikumar (1992)]. In this paper, I abstract from the life-cycle consumption/saving considerations and focus on how family background affects the child's college attendance decision and its implications on skill prices. In Appendix, I derive the optimal (constant) rate of intergenerational transfer in a standard overlapping-generations model where parents value their children's consumption and leave bequests to their children.

7. In a framework of multilevel CES production function that nests both capital-skill complementarity and skill-biased technical change as potential sources of the widening skill premium, McAdam and Willman (2017) show evidence that a production function where skilled and unskilled labor form a compound factor and the widening skill premium is explained by the faster skill-augmenting technical change fits the US data best. This is the form of the production function adopted in this paper.

8. The college completion rates in the data correspond to $\int_0^{\overline{\theta}} p_t(\theta)dG$ in the model (now allowed to be time varying). The construction of the data is discussed in detail in Section 3.1.

9. I thank an anonymous referee for making this point, that the wage gap can decrease following a relaxation of the budget constraint, when talent and initial wealth are positively correlated.

10. The derivative of the average talent of the unskilled with respect to x in the current setup is $\frac{d}{dx}\left[\frac{\frac{1}{2}\underline{\theta}+(\frac{1}{2}-\overline{p}x)\overline{\theta}}{1-\overline{p}x}\right] = -\frac{\frac{1}{2}\overline{p}(\overline{\theta}-\underline{\theta})}{[\frac{1}{2}+(1-\overline{p})x]^2}$. It is always strictly bigger in an absolute value than $\underline{W}'(x)$ in the static model in Section 2.1.

11. In Section 3.2, I use the discrete version of the dynamic system with each period equal to a year in the data. For example, the actual series of the efficiency unit of skilled labor in the simulation is

$$\psi_{st} = (1+\gamma_{\text{SBTC}})^t \frac{\int_0^{\overline{\theta}} \theta p(\theta)dG}{\int_0^{\overline{\theta}} p(\theta)dG}.$$

This is the notation I will adopt in Section 3.2.

12. The number 17.25% is obtained from 30% $\times$ 23/40, where 23 is the number of periods in my empirical analysis and 40 is the number of periods in Carneiro and Lee's.

REFERENCES

Acemoglu, D. (1998) Why do new technologies complement skills? Directed technical change and wage inequality. *Quaterly Journal of Economics* 113, 1055–1089.

Antras, P. (2004) Is the U.S. aggregate production function Cobb-Douglas? New estimates of the elasticity of substitution. *Contributions to Macroeconomics* 4, 1161–1162.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008) Trends in U.S. wage inequality: Revising the revisionists. *Review of Economics and Statistics* 90(2), 300–323.

Balart, P. (2016) The increase in college premium and the decline in low-skill wages: A signaling story. *Journal of Public Economic Theory* 18(3), 363–384.

Barro, R. J. (1974) Are government bonds net wealth? *Journal of Political Economy* 82(6), 1095–1117.

Baumol, W. J. and S. A. B. Blackman (1995) How to think about rising college costs. *Planning for Higher Education* 23(4), 1–7.

Bedard, K. (2001) Human capital versus signaling models: University access and high school dropouts. *Journal of Political Economy* 109(4), 749–775.

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

Blackburn, M. L., D. E. Bloom, and R. B. Freeman (1990) The declining economic position of less skilled American men. In G. Burtless (ed.), *A Future of Lousy Jobs? The Changing Structure of U.S. Wages*. Brookings Institution.

Bound, J. and G. Johnson (1992) Changes in the structure of wages in the 1980's: An evaluation of alternative explanations. *American Economics Review* 82(3), 371–392.

Cameron, S. V. and J. J. Heckman (2001) The dynamics of educational attainment for black, Hispanic, and white males. *Journal of Political Economy* 109(3), 455–499.

Card, D. and J. E. DiNardo (2002) Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics* 20(4), 733–783.

Card, D. and T. Lemieux (2001) Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *Quarterly Journal of Economics* 116(2), 705–746.

Carneiro, P. and J. J. Heckman (2002) The evidence on credit constraints in post-secondary schooling. *Economic Journal* 112(482), 705–734.

Carneiro, P. and S. Lee (2011) Trends in quality-adjusted skill premia in the United States, 1960–2000. *American Economic Review* 101(6), 2309–2349.

Eckstein, Z. and E. Nagypal (2004) The evolution of U.S. earnings inequality: 1961–2002. *Federal Reserve Bank of Minneapolis Quarterly Review* 28(2), 10–29.

Fang, H. (2006) Disentangling the college wage premium: Estimating a model with endogenous education choices. *International Economic Review* 47(4), 1151–1185.

Glomm, G. and B. Ravikumar (1992) Public versus private investment in human capital: Endogenous growth and income inequality. *Journal of Political Economy* 100(4), 818–834.

Guvenen, F. and B. Kuruscu (2012) Understanding the evolution of the US wage distribution: A theoretical analysis. *Journal of the European Economic Association* 10(3), 482–517.

He, H. (2012) What drives the skill premium: Technological change or demographic variation. *European Economic Review* 56(8), 1546–1572.

He, H. and Z. Liu (2008) Investment-specific technological change, skill accumulation, and wage inequality. *Review of Economic Dynamics* 11(2), 314–334.

Hendel, I., J. Shapiro, and P. Willen (2005) Educational opportunity and income inequality. *Journal of Public Economics* 89(5–6), 841–870.

Hendricks, L. and T. Schoellman (2014) Student abilities during the expansion of US education. *Journal of Monetary Economics* 63, 19–36.

Juhn, C., D. I. Kim, and F. Vella (2005) The expansion of college education in the United States: Is there evidence of declining cohort quality? *Economic Inquiry* 43(2), 303–315.

Katz, L. F. and K. M. Murphy (1992) Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics* 107(1), 35–78.

Krusell, P., L. E. Ohanian, J.-V. Rios-Rull, and G. L. Violante (2000) Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica* 68(5), 1029–1053.

Lochner, L. and A. Monge-Naranjo (2011) The nature of credit constraints and human capital. *American Economic Review* 101, 2487–2529.

Lochner, L. and A. Monge-Naranjo (2012) Credit constraints in education. *Annual Review of Economics* 4(1), 225–256.

Maoz, Y. D. and O. Moav (2004) Social stratification, capital-skill complementarity, and the nonmonotonic evolution of the education premium. *Macroeconomic Dynamics* 8, 295–309.

McAdam, P. and A. Willman (2017) Unraveling the skill premium. *Macroeconomic Dynamics*, 1–30. doi:10.1017/S1365100516000547.

Riley, J. G. (2001) Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature* 39(2), 432–478.

Spence, M. (1973) Job market signaling. *Quarterly Journal of Economics* 87(3), 355–374.

Stiglitz, J. E. (1975) The theory of 'screening', education, and the distribution of income. *American Economic Review* 65(3), 283–300.

Taber, C. R. (2001) The rising college premium in the eighties: Return to college or return to unobserved ability? *The Review of Economic Studies* 68(3), 665–691.

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

APPENDIX

Proof of Lemma 1. Let the net benefit of going to college $\Delta(\theta, k_0)$ be defined as the difference between the choice-specific flow value functions in (4) and (5), and we have

$$\begin{aligned}\Delta(\theta, k_0) &\equiv r\,(v^c(k;\theta,k_0) - v^{nc}(k;\theta,k_0))\\ &= \left(1-\phi+\phi\frac{dv}{dk}\right)[p(\theta)(\overline{W}-\underline{W}) - RQ] > 0 \Rightarrow p(\theta)(\overline{W}-\underline{W}) - RQ > 0.\end{aligned} \quad \textbf{(A.1)}$$

This immediately implies for all $\theta' > \theta$,

$$\Delta(\theta', k_0) = (1-\phi+\phi\frac{dv(k;\theta',k_0)}{dk})[p(\theta')(\overline{W}-\underline{W}) - RQ] > 0. \quad \textbf{(A.2)}$$

That is, independent of k, the agent (θ', k_0) would always prefer college as long as $k \geq Q$. ■

Proof of Lemma 2. If an agent indexed by (θ, k_0) can afford college at t, then it must be true that $k(t;\theta,k_0) \geq Q$. The law of motion of capital at the individual level is

$$\begin{aligned}\dot{k}(t;\theta,k_0) &= \phi\max\{p(\theta)\overline{W}(t) + (1-p(\theta))\underline{W}(t) + R(t)[k(t;\theta,k_0) - Q],\\ &\underline{W}(t) + R(t)k(t;\theta,k_0)\} \geq \phi[\underline{W}(t) + R(t)k(t;\theta,k_0)] > 0.\end{aligned} \quad \textbf{(A.3)}$$

Since the capital stock at the individual level always grows at a positive rate (under any positive factor prices), once $k(t;\theta,k_0) \geq Q$, then $k(t';\theta,k_0) \geq Q$ for all t' greater than t. ■

Proof of Proposition 1. The sufficient and necessary condition of the existence of a wealth-separating equilibrium is to guarantee that agents of all talent find attending college attractive so that financial resources become the only hurdle to college enrollment. By *Lemma 1*, it is enough to show that schooling is an optimal choice for the least talented at all times given the equilibrium prices. The least talented prefers college if

$$p(0)[\overline{W}(t) - \underline{W}(t)] \geq R(t)Q. \quad \textbf{(A.4)}$$

It is necessary that $\overline{W}(t) > \underline{W}(t)$ or $\pi(t) > 1$ for all t. Since in the equilibrium $\pi(t)$ always increases, it is enough that

$$\pi(0) = \frac{1-\lambda}{\lambda}\left(\frac{h_u(0)}{h_s(0)}\right)^{1-\rho}\left(\frac{\psi_s(0)}{\psi_u(0)}\right)^{\rho} > 1 \Leftrightarrow \lambda < \frac{1}{\left(\frac{h_s(0)}{h_u(0)}\right)^{1-\rho}\left(\frac{\psi_u(0)}{\psi_s(0)}\right)^{\rho}+1}, \quad \textbf{(A.5)}$$

where

$$\frac{h_s(0)}{h_u(0)} = \frac{x(0)\int_0^{\overline{\theta}} p(\theta)dG}{1 - x(0)\int_0^{\overline{\theta}} p(\theta)dG}$$

$$\text{and}\quad \frac{\psi_u(0)}{\psi_s(0)} = \frac{\int_0^{\overline{\theta}}\theta dG - x(0)\int_0^{\overline{\theta}}\theta p(\theta)dG}{1 - x(0)\int_0^{\overline{\theta}} p(\theta)dG}\,\frac{\int_0^{\overline{\theta}} p(\theta)dG}{\int_0^{\overline{\theta}}\theta p(\theta)dG}. \quad \textbf{(A.6)}$$

Notice from (11) the growth of the skill premium depends on ρ. In particular, ρ needs to be high enough so the relative efficiency effect from signaling can overcome the relative quantity effect resulting from the declining marginal products:

$$\frac{d\ln\pi}{dx} = (1-\rho)\frac{d}{dx}\ln\frac{h_u}{h_s} + \rho\frac{d}{dx}\ln\frac{\psi_s}{\psi_u}$$

$$= (1-\rho)\frac{d}{dx}\left[\ln\frac{1-x\int_0^{\overline{\theta}}p(\theta)dG}{x\int_0^{\overline{\theta}}p(\theta)dG}\right] + \rho\frac{d}{dx}\left[\ln\frac{1-x\int_0^{\overline{\theta}}p(\theta)dG}{\int_0^{\overline{\theta}}\theta dG - x\int_0^{\overline{\theta}}\theta p(\theta)dG}\right] \quad \textbf{(A.7)}$$

$$= (1-\rho)\frac{-1}{x(1-x\int_0^{\overline{\theta}}p(\theta)dG)} + \rho\frac{\int_0^{\overline{\theta}}\theta p(\theta)dG - \int_0^{\overline{\theta}}p(\theta)dG\int_0^{\overline{\theta}}\theta dG}{\left(1-x\int_0^{\overline{\theta}}p(\theta)dG\right)\left(\int_0^{\overline{\theta}}\theta dG - x\int_0^{\overline{\theta}}\theta p(\theta)dG\right)} \geq 0$$

$$\Leftrightarrow \frac{1}{1-\rho} \geq 1 + \frac{\int_0^{\overline{\theta}}\theta dG - x\int_0^{\overline{\theta}}\theta p(\theta)dG}{x\left(\int_0^{\overline{\theta}}\theta p(\theta)dG - \int_0^{\overline{\theta}}p(\theta)dG\int_0^{\overline{\theta}}\theta dG\right)}. \quad \textbf{(A.8)}$$

The elasticity of substitution between skilled and unskilled labor must necessarily be greater than 1 in order for the relative efficiency effect to dominate. The RHS of the above inequality is clearly decreasing in x; therefore, a sufficient condition to guarantee a positive growth in the skill premium is

$$\frac{1}{1-\rho} \geq 1 + \frac{\int_0^{\overline{\theta}}\theta dG - x(0)\int_0^{\overline{\theta}}\theta p(\theta)dG}{x(0)\left(\int_0^{\overline{\theta}}\theta p(\theta)dG - \int_0^{\overline{\theta}}p(\theta)dG\int_0^{\overline{\theta}}\theta dG\right)} \quad \textbf{(A.9)}$$

$$\Leftrightarrow \rho \geq \frac{\int_0^{\overline{\theta}}\theta dG - x(0)\int_0^{\overline{\theta}}\theta p(\theta)dG}{\int_0^{\overline{\theta}}\theta dG - x(0)\int_0^{\overline{\theta}}p(\theta)dG\int_0^{\overline{\theta}}\theta dG}. \quad \textbf{(A.10)}$$

Substituting the factor prices into (A.4), it becomes equivalent to

$$p(0)(1-\mu)[\lambda u^{\rho}+(1-\lambda)s^{\rho}]^{\frac{\sigma}{\rho}-1}[(1-\lambda)s^{\rho-1}\psi_s-\lambda u^{\rho-1}\psi_u] \geq \mu(K-xQ)^{\sigma-1}Q. \quad \textbf{(A.11)}$$

Let the LHS of the above inequality be denoted as $\Phi(x)$ and the RHS be denoted as $\Psi(x, Q)$. Consider

$$\Phi(x) = \Psi(x, Q). \quad \textbf{(A.12)}$$

For a given value of x, $\Phi(x)$ is a strictly positive number, and $\Psi(x, Q)$ is strictly increasing in Q, with $\Psi(x, 0) = 0$ and $\Psi(x, Q) \to +\infty$, as $Q \to \frac{K}{x}-$: $\frac{d}{dQ}[\mu(K-xQ)^{\sigma-1}Q] = \mu(K-xQ)^{\sigma-2}(K-\sigma xQ) > 0$. Therefore, there exists a unique $\widehat{q} \in (0, \frac{K}{x})$ such that (A.12) holds at $\widehat{q}$ for a given x. Write it as $\widehat{q}(x)$. Note that by the continuity of $\Phi(x)$ and $\Psi(x)$, $\widehat{q}(x)$ is continuous in x and strictly positive. Define $\widehat{Q} = \min_{x\in[x(0),1]}\widehat{q}(x)$, which exists. Then, for all $Q \leq \widehat{Q}$, the inequality (A.4) holds.

Note that a dynasty starts to send agents to college at t if its initial capital endowment $\widehat{k}_0(t)$ satisfies

$$\widehat{k}_0(t) + \int_0^t \dot{k}(s;\theta,\widehat{k}_0(t))ds = Q, \quad \textbf{(A.13)}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

where the evolution of $k(t;\theta,k_0)$ follows $\dot{k}(t;\theta,k_0) = \phi[R(t)k(t;\theta,k_0) + \underline{W}(t)]$. Take derivative of (A.13) with respect to t,

$$\widehat{k}_0'(t) = -\phi[R(t)k(t;\theta,\widehat{k}_0(t)) + \underline{W}(t)] < 0. \quad \textbf{(A.14)}$$

At time t, the faction of agents that go to college is $x(t) = 1 - F(\widehat{k}_0(t))$, which is clearly increasing in t. The equilibrium path can be completely characterized by the dynamic system in the aggregate capital, $K(t)$, and the cutoff wealth level, $\widehat{k}_0(t)$, given in (24). ■

Proof of Proposition 2. Step 1: Transformation. Let $\widehat{p}(\theta) = \overline{\theta}p(\theta)g(\theta)$, which necessarily satisfies $\widehat{p}(\theta) \geq 0$ and $\eta\overline{\theta} \leq \int_0^{\overline{\theta}} \widehat{p}(\theta)d\theta \leq \overline{\theta}$. Let $\int_0^{\overline{\theta}} \theta dG \equiv \mu_\theta$. Then, $\int_0^{\overline{\theta}} \theta\widehat{p}d\theta / \int_0^{\overline{\theta}} \widehat{p}(\theta)d\theta \geq \xi\mu_\theta$. This problem is equivalent to a two-step maximization problem. Given μ_θ, first solve

$$\sup_{\widehat{p}(\theta)}(\text{or inf})\dot{x}\overline{\theta}\frac{\int_0^{\overline{\theta}}\theta\widehat{p}d\theta - \mu_\theta\int_0^{\overline{\theta}}\widehat{p}d\theta}{\left(\overline{\theta} - x\int_0^{\overline{\theta}}\widehat{p}d\theta\right)\left(\overline{\theta}\mu_\theta - x\int_0^{\overline{\theta}}\theta\widehat{p}d\theta\right)} \quad \textbf{(A.15)}$$

$$\text{s.t.}\begin{cases}\widehat{p}(\theta) \geq 0\\ \eta\overline{\theta} \leq \int_0^{\overline{\theta}}\widehat{p}(\theta)d\theta \leq \overline{\theta}\\ \xi\mu_\theta\int_0^{\overline{\theta}}\widehat{p}(\theta)d\theta \leq \int_0^{\overline{\theta}}\theta\widehat{p}d\theta \leq \mu_\theta\overline{\theta}\end{cases}. \quad \textbf{(A.16)}$$

Then, optimize over all possible μ_θ.

Step 2: Change of variables. Let $y(\theta) = \int_0^\theta \widehat{p}(v)dv$. Integration by part gives $\int_0^{\overline{\theta}}\theta\widehat{p}d\theta = \int_0^{\overline{\theta}}\theta y'(\theta)d\theta = \overline{\theta}y(\overline{\theta}) - \int_0^{\overline{\theta}}y(\theta)d\theta$. The problem can be rewritten as

$$\sup_{\substack{y(\overline{\theta}),\\ \int_0^{\overline{\theta}}y(\theta)d\theta}}(\text{or inf})\dot{x}\overline{\theta}\frac{(\overline{\theta} - \mu_\theta)y(\overline{\theta}) - \int_0^{\overline{\theta}}y(\theta)d\theta}{(\overline{\theta} - xy(\overline{\theta}))\left[x\int_0^{\overline{\theta}}y(\theta)d\theta + \overline{\theta}(\mu_\theta - xy(\overline{\theta}))\right]} \quad \textbf{(A.17)}$$

$$\text{s.t.}\begin{cases}\eta\overline{\theta} \leq y(\overline{\theta}) \leq \overline{\theta}\\ y'(\theta) \geq 0\\ \max\{0, \overline{\theta}(y(\overline{\theta}) - \mu_\theta)\} \leq \int_0^{\overline{\theta}}y(\theta)d\theta \leq (\overline{\theta} - \xi\mu_\theta)y(\overline{\theta})\end{cases}. \quad \textbf{(A.18)}$$

Step 3: Optimization. First, $y(\overline{\theta})$ and $\int_0^{\overline{\theta}}y(\theta)d\theta$ can take values independently. Second, the objective is increasing in $y(\overline{\theta})$, but decreasing in $\int_0^{\overline{\theta}}y(\theta)d\theta$. However, note that the value of $y(\overline{\theta})$ will affect the boundaries of the values that $\int_0^{\overline{\theta}}y(\theta)d\theta$ can take due to (A.18).

Consider the *sup* problem first. If $y(\overline{\theta}) \leq \mu_\theta$, then the optimal values are $y(\overline{\theta}) = \mu_\theta$ and $\int_0^{\overline{\theta}}y(\theta)d\theta = 0$. If $y(\overline{\theta}) \geq \mu_\theta$, then at the optimum, for a given $y(\overline{\theta})$, set $\int_0^{\overline{\theta}}y(\theta)d\theta = \overline{\theta}(y(\overline{\theta}) - \mu_\theta)$. Substituting this equation into the objective function: $\sup_{y(\overline{\theta})}\dot{x}\frac{\overline{\theta}-y(\overline{\theta})}{(\overline{\theta}-xy(\overline{\theta}))(1-x)}$, decreasing in $y(\overline{\theta})$. Hence, at the optimum, $y(\overline{\theta}) = \mu_\theta$ and $\int_0^{\overline{\theta}}y(\theta)d\theta = 0$. In both cases, $\sup(g_{\psi_s} - g_{\psi_u}) = \dot{x}\frac{\overline{\theta}-\mu_\theta}{(\overline{\theta}-x\mu_\theta)(1-x)}$. Now maximize the above with respect to μ_θ: $\sup(g_{\psi_s} - g_{\psi_u}) = \frac{\dot{x}}{1-x} = -g_{1-x}$, as $\mu_\theta \to 0$.

Now consider the inf problem. For a given $y(\overline{\theta})$, set $\int_0^{\overline{\theta}} y(\theta)d\theta = (\overline{\theta} - \xi\mu_\theta)y(\overline{\theta})$. Substituting this equation into the objective function: $\inf_{y(\overline{\theta})} \dot{x}\overline{\theta}\frac{y(\overline{\theta})(\xi-1)}{(\overline{\theta}-xy(\overline{\theta}))(\overline{\theta}-x\xi y(\overline{\theta}))}$, increasing in $y(\overline{\theta})$. Hence, setting $y(\overline{\theta})$ to $\eta\overline{\theta}$, I have $\inf(g_{\psi_s} - g_{\psi_u}) = \frac{\eta(\xi-1)\dot{x}}{(1-x\eta)(1-x\eta\xi)}$. ■

A.1. OVERLAPPING GENERATIONS MODEL WITH BEQUEST

Consider a stylized, small, open-economy, overlapping generations model with constant (unit) population. Each generation lives for two periods and has identical preferences à la Barro (1974). Generation t, endowed with k_t from its previous generation, consumes c_t^y when young, consumes c_{t+1}^o when old, and leave k_{t+1} to the next generation when old. His life-time utility is given by

$$U(c_t^y, c_{t+1}^o, U_{t+1}^*) = \ln c_t^y + \beta \ln c_{t+1}^o + \phi U_{t+1}^*, \quad \textbf{(A.19)}$$

where U_{t+1}^* is the life-time utility achieved by his immediate offspring. He maximizes his life-time utility choosing own consumption c_t^y and c_{t+1}^o, and his bequest to the next generation k_{t+1}, subject to the budget constraint:

$$k_{t+1} = R(k_t - c_t^y) - c_{t+1}^o. \quad \textbf{(A.20)}$$

Denoting the stationary value function as $V(k)$, the Bellman equation of this recursive problem is

$$V(k) = \max_{c^y, c^o, k'} \{\ln c^y + \beta \ln c^o + \phi V(k')\}, \quad \textbf{(A.21)}$$

subject to (A.20). Guess the value function takes the functional form $V(k) = a + b\ln(k)$. The FOC with respect to c^y implies

$$\frac{1}{c^y} = \phi R \frac{b}{k'} \Rightarrow c^y = \frac{k'}{\phi R b}. \quad \textbf{(A.22)}$$

The FOC with respect to c^o implies

$$\frac{\beta}{c^o} = \phi \frac{b}{k'} \Rightarrow c^o = \frac{\beta k'}{\phi b}. \quad \textbf{(A.23)}$$

From the budget constraint

$$k' = R(k - \frac{k'}{\phi R b}) - \frac{\beta k'}{\phi b} \Rightarrow k' = \frac{\phi R b}{\phi b + 1 + \beta} k. \quad \textbf{(A.24)}$$

Plug (A.22)–(A.24) to the Bellman equation:

$$\begin{aligned} a + b\ln k &= \ln\left(\frac{k'}{\phi R b}\right) + \beta \ln\left(\frac{\beta k'}{\phi b}\right) + \phi\left[a + b\ln\left(k'\right)\right] \\ &= \ln\left(\frac{k}{\phi b + 1 + \beta}\right) + \beta \ln\left(\frac{\beta R k}{\phi b + 1 + \beta}\right) + \phi\left[a + b\ln\left(\frac{\phi R b k}{\phi b + 1 + \beta}\right)\right] \quad \textbf{(A.25)} \end{aligned}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

Equating the coefficients, I can solve for a and b. The optimal consumption and intergenerational transfer plans are given by

$$k' = \phi R k, \tag{A.26}$$

$$c^y = \frac{1-\phi}{1+\beta} k, \tag{A.27}$$

$$c^o = \frac{\beta R(1-\phi)}{1+\beta} k. \tag{A.28}$$

In this highly stylized model of bequest, the parent generation always bequeathes a constant fraction of their wealth to the next immediate generation. Compare (A.26) to (4) and (5) in the paper, the rate of intergenerational transfer ϕ measures the degree of altruism in the parent's utility function.

A.2. OTHER TYPES OF EQUILIBRIA

In general, there exist other types of equilibria where the schooling decision not only depends on the capital holding but also the ability. In such a scenario, Lemma 1 continues to hold. Therefore, one could imagine a separating equilibrium where agents whose talent is above some threshold and who are not constrained attend college. In this section, I establish the existence of an equilibrium that is separating both in initial wealth and in talent, under the special case where skilled and unskilled labor are perfect substitutes (i.e., $\rho = 1$). I follow the notation in the paper, except that the fraction of the unconstrained agents is no longer equal to the rate of enrollment. I use $z(t)$ for the former and keep $x(t)$ for the latter. Suppose

Assumption 1.

$$p(0) = 0.$$

Assumption 2.

$$p(E(\theta))\left[(1-\lambda)\frac{\int_{E(\theta)}^{\overline{\theta}} \theta p dG}{\int_{E(\theta)}^{\overline{\theta}} p dG} - \lambda \frac{E(\theta) - z(0)\int_{E(\theta)}^{\overline{\theta}} \theta p dG}{1 - z(0)\int_{E(\theta)}^{\overline{\theta}} p dG}\right] E(\theta)^{\sigma-1}$$
$$\geq \frac{\mu}{1-\mu}\left[k(0) - (1 - G(E(\theta)))Q\right]^{\sigma-1} Q,$$

where $k(0) = \int_0^{\overline{k}_0} k_0 dF$ and $z(0) = 1 - F(Q)$.

Assumption 3.

$$E\left[p(\theta)|\theta \geq E(\theta)\right] \leq \frac{\sqrt{1-\lambda}}{\sqrt{\lambda} + \sqrt{1-\lambda}}.$$

PROPOSITION 3. *Let $\rho = 1$. Under Assumptions 1, 2, and 3, there is an equilibrium separating in both initial wealth and talent so that all agents whose initial capital is above $\widehat{k}_0(t)$ and whose talent is above $\widehat{\theta}(t)$ go to college.*

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

Proof. Let the fraction of the unconstrained agents be denoted $z(t) = 1 - F(\widehat{k}_0(t))$. Since Lemma 1 continues to hold, the cutoff talent $\widehat{\theta}(t)$ must satisfy

$$p(\widehat{\theta}(t))[\overline{W}(t) - \underline{W}(t)] = R(t)Q. \tag{A.29}$$

Under the factor prices when $\rho = 1$, the above condition is equivalent to

$$p(\widehat{\theta}(t))\left[(1-\lambda)\frac{\int_{\widehat{\theta}(t)}^{\overline{\theta}} \theta p dG}{\int_{\widehat{\theta}(t)}^{\overline{\theta}} p dG} - \lambda\frac{E(\theta) - z(t)\int_{\widehat{\theta}(t)}^{\overline{\theta}} \theta p dG}{1 - z(t)\int_{\widehat{\theta}(t)}^{\overline{\theta}} p dG}\right] E(\theta)^{\sigma-1}$$
$$= \frac{\mu}{1-\mu}\left[k(t) - z(t)\left(1 - G(\widehat{\theta}(t))\right) Q\right]^{\sigma-1} Q. \tag{A.30}$$

Denote the RHS of the above equation as RHS$(\widehat{\theta})$ and the LHS as LHS$(\widehat{\theta})$. Clearly, RHS$(\widehat{\theta})$ is decreasing in $\widehat{\theta}$. LHS$(\widehat{\theta})$ is increasing in $\widehat{\theta}$ if (a) $\widehat{\theta} < E(\theta)$ and (b) $z(t)\int_{\widehat{\theta}}^{\overline{\theta}} p dG < \frac{\sqrt{1-\lambda}}{\sqrt{\lambda}+\sqrt{1-\lambda}}$. To see this,

$$\frac{d}{d\widehat{\theta}}\left[(1-\lambda)\frac{\int_{\widehat{\theta}}^{\overline{\theta}} \theta p dG}{\int_{\widehat{\theta}}^{\overline{\theta}} p dG} - \lambda\frac{E(\theta) - z(t)\int_{\widehat{\theta}}^{\overline{\theta}} \theta p dG}{1 - z(t)\int_{\widehat{\theta}}^{\overline{\theta}} p dG}\right]$$
$$= p(\widehat{\theta})g(\widehat{\theta})\left\{\lambda z(t)\frac{E(\theta) - \widehat{\theta}}{\left[1 - z(t)\int_{\widehat{\theta}}^{\overline{\theta}} p dG\right]^2}\right.$$
$$\left. + \int_{\widehat{\theta}}^{\overline{\theta}}(\theta - \widehat{\theta}) p dG\left[\frac{1-\lambda}{\left(\int_{\widehat{\theta}}^{\overline{\theta}} p dG\right)^2} - \frac{\lambda z(t)^2}{\left(1 - z(t)\int_{\widehat{\theta}}^{\overline{\theta}} p dG\right)^2}\right]\right\}$$
$$> 0 \text{ under (a) and (b).} \tag{A.31}$$

Under Assumption 1, I have LHS$(0) = 0 <$ RHS$(0) = \frac{\mu}{1-\mu}[k(t) - z(t)Q]^{\sigma-1} Q$. Under Assumption 2, I have LHS$(E(\theta)) >$ RHS$(E(\theta))$:

$$\text{LHS}(E(\theta)) = p(E(\theta))\left[(1-\lambda)\frac{\int_{E(\theta)}^{\overline{\theta}} \theta p dG}{\int_{E(\theta)}^{\overline{\theta}} p dG} - \lambda\frac{E(\theta) - z(t)\int_{E(\theta)}^{\overline{\theta}} \theta p dG}{1 - z(t)\int_{E(\theta)}^{\overline{\theta}} p dG}\right] E(\theta)^{\sigma-1}$$
$$> p(E(\theta))\left[(1-\lambda)\frac{\int_{E(\theta)}^{\overline{\theta}} \theta p dG}{\int_{E(\theta)}^{\overline{\theta}} p dG} - \lambda\frac{E(\theta) - z(0)\int_{E(\theta)}^{\overline{\theta}} \theta p dG}{1 - z(0)\int_{E(\theta)}^{\overline{\theta}} p dG}\right] E(\theta)^{\sigma-1}$$
$$\geq \frac{\mu}{1-\mu}[k(0) - (1 - G(E(\theta)))\, Q]^{\sigma-1} Q$$
$$> \frac{\mu}{1-\mu}[k(t) - z(t)\,(1 - G(E(\theta)))\, Q]^{\sigma-1} Q = \text{RHS}(E(\theta)). \tag{A.32}$$

The first inequality follows since $\frac{E(\theta) - z\int_{E(\theta)}^{\overline{\theta}} \theta p dG}{1 - z\int_{E(\theta)}^{\overline{\theta}} p dG}$ decreases in z. The second is given by Assumption 2 and the last inequality is obvious. Therefore, the solution $\widehat{\theta}(t)$ exists and

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

must lie in the open interval $(0, E(\theta))$, which together with Assumption 3 implies (b):

$$z(t)\int_{\widehat{\theta}}^{\overline{\theta}} pdG = z(t)[1-G(\widehat{\theta})]\frac{\int_{\widehat{\theta}}^{\overline{\theta}} pdG}{1-G(\widehat{\theta})} = z(t)[1-G(\widehat{\theta})]E[p(\theta)|\theta \geq E(\theta)]$$

$$< [1-G(\widehat{\theta})]E[p(\theta)|\theta \geq E(\theta)] < E[p(\theta)|\theta \geq E(\theta)] \leq \frac{\sqrt{1-\lambda}}{\sqrt{\lambda}+\sqrt{1-\lambda}}. \quad \textbf{(A.33)}$$

■

In this economy, the quantity and efficiency units of skilled and unskilled labor are given by

$$E_t(\theta|s) = \frac{\int_{\widehat{\theta}(t)}^{\overline{\theta}} \theta p(\theta)dG}{\int_{\widehat{\theta}(t)}^{\overline{\theta}} p(\theta)dG} \quad \textbf{(A.34)}$$

$$E_t(\theta|u) = \frac{E(\theta) - z(t)\int_{\widehat{\theta}(t)}^{\overline{\theta}} \theta p(\theta)dG}{1 - z(t)\int_{\widehat{\theta}(t)}^{\overline{\theta}} p(\theta)dG}; \quad \textbf{(A.35)}$$

$$h_s(t) = z(t)\int_{\widehat{\theta}(t)}^{\overline{\theta}} p(\theta)dG; \quad \textbf{(A.36)}$$

$$h_u(t) = 1 - z(t)\int_{\widehat{\theta}(t)}^{\overline{\theta}} p(\theta)dG. \quad \textbf{(A.37)}$$

This separating equilibrium is characterized by the system below. Let the enrollment rate be denoted $x(t) = z(t)(1-G(\widehat{\theta}))$.

$$\begin{cases} \dot{K}(t) = \phi Y\left(K(t) - x(t)Q, E(\theta) - z(t)\int_{\widehat{\theta}(t)}^{\overline{\theta}} \theta p(\theta)dG, z(t)\int_{\widehat{\theta}(t)}^{\overline{\theta}} \theta p(\theta)dG\right) \\ \dot{\widehat{k}}_0(t) = -\phi\left[R(t)Q + \underline{W}(t)\right] \end{cases} \quad \textbf{(A.38)}$$

and the talent threshold $\widehat{\theta}(t)$ satisfies (A.30).

In this equilibrium, as dynasties accumulate wealth, the fraction of the unconstrained dynasties, $z(t)$, increases over time, and the talent type $\widehat{\theta}(t)$ that is indifferent between attending college and not evolves within the interval $(0, E(\theta))$. However, along the equilibrium path, the talent threshold, the enrollment rates and the college premium are not necessarily monotone. The following corollary gives the condition under which the skill premium decreases in the talent threshold. One can conceive an equilibrium where dynasties are less constrained, the marginal college goer less talented, the enrollment increases and the skill premium increases over time.

COROLLARY. The skill premium $\ln\frac{\overline{W}(t)}{\underline{W}(t)}$ increases in the fraction of the unconstrained dynasties, $z(t)$. It decreases in the threshold talent $\widehat{\theta}(t)$ if and only if $\widehat{\theta} > \frac{E(\theta|u)E(\theta|s)}{E(\theta)}$.

Proof. The log wage gap is

$$\ln\frac{\overline{W}}{\underline{W}} = \ln\frac{\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG}{\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG} - \ln\frac{E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG}{1 - z\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG} + \ln\frac{1-\lambda}{\lambda}. \quad \textbf{(A.39)}$$

Take the derivative of the above with respect to z,

$$\frac{d}{dz}\left(\ln\frac{\overline{W}}{\underline{W}}\right) = \frac{\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG - E(\theta)\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG}{\left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG\right)\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG\right)}$$

$$= \frac{\int_{\widehat{\theta}}^{\overline{\theta}} [\theta - E(\theta)]p(\theta)dG}{\left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG\right)\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG\right)}$$

$$= \frac{\int_{\widehat{\theta}}^{E(\theta)} [\theta - E(\theta)]p(\theta)dG + \int_{E(\theta)}^{\overline{\theta}} [\theta - E(\theta)]p(\theta)dG}{\left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG\right)\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG\right)}$$

$$> \frac{p(E(\theta))\int_{\widehat{\theta}}^{E(\theta)} [\theta - E(\theta)]dG + p(E(\theta))\int_{E(\theta)}^{\overline{\theta}} [\theta - E(\theta)]dG}{\left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG\right)\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG\right)}$$

$$= \frac{p(E(\theta))\int_{\widehat{\theta}}^{\overline{\theta}} [\theta - E(\theta)]dG}{\left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG\right)\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG\right)}. \quad \textbf{(A.40)}$$

Note that

$$\frac{d}{d\widehat{\theta}}\int_{\widehat{\theta}}^{\overline{\theta}} [\theta - E(\theta)]dG = -[\widehat{\theta} - E(\theta)]g(\widehat{\theta}) > 0. \quad \textbf{(A.41)}$$

Therefore,

$$\frac{d}{dz}\left(\ln\frac{\overline{W}}{\underline{W}}\right) > \frac{p(E(\theta))\int_{0}^{\overline{\theta}} [\theta - E(\theta)]dG}{\left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta p(\theta)dG\right)\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} p(\theta)dG\right)} = 0. \quad \textbf{(A.42)}$$

Take derivative of (A.43) with respect to $\widehat{\theta}$,

$$\frac{d}{d\widehat{\theta}}\left(\ln\frac{\overline{W}}{\underline{W}}\right) < 0 \Leftrightarrow \quad \frac{\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG - \widehat{\theta}\int_{\widehat{\theta}}^{\overline{\theta}} pdG}{\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG \int_{\widehat{\theta}}^{\overline{\theta}} pdG}$$

$$< z\frac{\widehat{\theta}\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} pdG\right) - \left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG\right)}{\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} pdG\right)\left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG\right)}$$

$$\Leftrightarrow \frac{1}{\int_{\widehat{\theta}}^{\overline{\theta}} pdG} - \frac{\widehat{\theta}}{\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG} < \frac{z\widehat{\theta}}{E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG} - \frac{z}{1 - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG}$$

$$\Leftrightarrow \frac{1}{\int_{\widehat{\theta}}^{\overline{\theta}} pdG} + \frac{z}{1 - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG} < \widehat{\theta}\left[\frac{1}{\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG} + \frac{z}{E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG}\right]$$

$$\Leftrightarrow \frac{1}{\int_{\widehat{\theta}}^{\overline{\theta}} pdG\left(1 - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG\right)} < \widehat{\theta}\frac{E(\theta)}{\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG\left(E(\theta) - z\int_{\widehat{\theta}}^{\overline{\theta}} \theta pdG\right)}$$

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press

$$\Leftrightarrow \frac{\int_{\widehat{\theta}}^{\overline{\theta}} \theta p dG \left(E(\theta) - z \int_{\widehat{\theta}}^{\overline{\theta}} \theta p dG \right)}{\int_{\widehat{\theta}}^{\overline{\theta}} p dG \left(1 - z \int_{\widehat{\theta}}^{\overline{\theta}} \theta p dG \right)} < \widehat{\theta} E(\theta)$$

$$\Leftrightarrow E(\theta|s)E(\theta|u) < \widehat{\theta} E(\theta). \tag{A.43}$$

■

As in the pooling equilibrium, the relaxation of the budget constraint tends to increase the skill premium. However, if the increase in enrollment is brought by a decrease in the talent threshold, the effect of increased college attendance on the skill premium is ambiguous. The corollary suggests that, if the marginal college-goer is still relatively talented, then allowing this person to attend college lowers the average talent of the high school graduates more than it lowers the average talent of the college graduates.

https://doi.org/10.1017/S1365100517000232 Published online by Cambridge University Press