Some Thoughts on Job Market Signaling with Human Capital

By Yangfan Wang*

This paper extends the original Spence signaling model by incorporating human capital formation from education. With this feature of human capital formation, education is no longer a pure cost and allows for worker optimization. In short-run equilibria in a two-sector job market, each worker has an optimal education level and a range of education for which they will choose high-paying jobs. If firms can observe all worker characteristics, they will have different minimum education requirements for workers of different ability levels. In contrast, under imperfect information, the firm can only observe worker education. The imperfect information may create incentives for some workers to deviate from their education levels in the perfect information equilibrium. The exact equilibrium states depend on specific interactions between the firm production function and the worker education cost function. However, it is possible that in the signaling equilibrium, workers with middle-range ability levels receive excessive education compared to the equilibrium under perfect information. This special case of this more general model is consistent with the original Spence model while allowing for more flexibility. On the other hand, in certain circumstances, the human capital effect of education may alleviate the inefficiency of the signalling equilibrium. This model also suggests that standard empirical estimation of the return to education tends to overestimate it, but a quantile regression or parametric mean regression at tails could address this issue.

I. Introduction

Although the rise of education level is often credited as an improvement of the labor force, the problem of overeducation also emerges. Overeducation refers to the phenomenon that workers obtain an education level higher than that required by their occupation. This phenomenon prevails in both developed and developing countries (see McGuinness, 2006 for a review). More recently, according to European Commission (2017), more than 20% of young workers in EU are overqualified for their jobs. Similar findings are also discovered in developing countries. Overeducation and credential inflation are found among 29%-43% of

* Graduate Student at Duke University, Email: yangfan.wang@duke.edu. I appreciate Dr. Charles M. Becker and Meghna Baskar for valuable comments and feedback on previous drafts. Special thanks to Alicia Gong, without whose persistent support this paper would have been impossible. I would also like to thank Yiheng Hao and Inkoo Back for inspiring discussions. Any error remains my own.

the workforce in Latin America (Castro et al., 2022), and the portion of overeducated and undereducated workers are estimated to be one quarter (Marioni, 2021).

The prevalence of overeducation suggests the inefficiency of the labor market and harms workers. The direct harm appears in the form of heavier cost of education: Bleemer et al. (2021) show that the average public college tuition per school year across states increased by 81% from 2001 to 2009 in the United States and ranged from $5,010 to $25,231 in 2013. Along with the rise in the education cost on the per-year basis, overinvestment in the level of education increases the overall cost of education. Using the 2013 public college tuition rate, an excessive four-year public college degree translates to additional financial cost of $20,040 − $100,924[1]. According to Groot and Maassen van den Brink (2000), the average overeducation rate in the United States was $26.3% during 1980-2000. Considering the size of labor force and rising tuition, the overall financial burden of overeducation is huge. Meanwhile, the indirect harm takes various forms, including wage penalties (Cultrera et al., 2022), lower occupation transitions (Baert, Cockx and Verhaest, 2013), and lower job satisfaction (Tsang, Rumberger and Levin, 1991).

Given the proliferation of overeducation, the reason behind it is worth studying. The classical explanation often follows the job market signaling model by Spence (1973). The original work of Spence takes an extreme view of education purely as a signal for ability and concludes with an equilibrium with overeducation. It contrasts with the human capital theory developed by Becker (1964), which assumes education improves abilities in production.[2] While there is plenty of empirical evidence confirming the existence of signaling effect of education (e.g., see Barrera-Osorio and Bayona-Rodríguez, 2019; Naven and Whalen, 2022; Garcia-Mainar and Montuenga, 2019), the body of literature does not completely negate the validity of human capital theory. Spence focused on the signaling aspect of education, and incorporation of human capital formation in the Spence model can generate a richer set of results.

This paper aims at extending the Spence (1973) signaling model by incorporat-

[1] Note that this statistic only surveys the tuition of public universities, while there are many more expensive private universities. For instance, the total estimated bill expenses at Duke University is $85,238 for undergraduate students and $93,321.50 for master's students during the 2023-2024 academic year, according to the Karsh Office of Undergraduate Financial Support, https://financialaid.duke.edu/how-aid-calculated/cost-attendance/ and the Graduate School, https://gradschool.duke.edu/financial-support/cost-attend/. A four-year undergraduate degree at Duke University then translates into financial cost of $340,952, and a two-year master's degree translates into $186,643.

[2] There are also other explanations for investment in education. For instance, positive assortative mating is also a reason for investing in education (see Chiappori, Iyigun and Weiss (2009)). Higher educated individuals tend to find spouses of higher quality and raise the quality of their children. However, since in our discussion we do not introduce the complexity of heterogeneous gender decisions or intergenerational utilities, positive assortative mating is less of our focus.

ing the effect on productivity improvement of education. Specifically, we focus on new entrants into the labor market and assume education affects job market outcomes through two channels: signaling of worker ability and formation of human capital. We follow Kalemli-Ozcan, Ryder and Weil (2000) and further assume workers earn wages proportional to their human capital. Hence, together with the rising cost of education, the workers solve for the maximization of their utility and a signaling game. Meanwhile, in contrast to the original Spence model, education does not serve as pure signals anymore, but enters the value function of firms. On the one hand, firms desire better educated workers, as it improves human capital; on the other hand, better educated workers are more costly. As a result, the firm also balances its costs and benefits. We define overeducation as the scenario when a worker acquires an education level higher than the level that would bring her the maximal net payoff. Under reasonable circumstances, the model still arrives at the overeducation scenario and provides flexibility in explaining varying degrees of overeducation across industries and regions.

One important result of this model is that in the signaling equilibrium, if overeducation occurs, it tends to occur only within a range of workers with middle ability and education. Since we assume that the cost of education increases with education and decreases with worker ability, it increases more rapidly for worker with lower ability. Therefore, the wage premium of more education for low-ability workers cannot compensate for the cost of education beyond a certain education level. Hence, only workers with higher ability can afford to acquire more education. This positive relationship between education and ability suggests that workers with too low ability cannot pretend to have higher ability by acquiring excessive education, since the wage premium cannot compensate for their high cost of education. Similarly, workers with very high ability do not need to signal their high ability by over-investing in education, since their education level is already high enough so that workers with lower abilities cannot afford to pretend to be them.

The rest of this paper is structured as follows. Section 2 describes the basic model setup. Section 3 describes the worker's optimization problem and derives a series of properties related to it. Section 4 describes the firm's objectives and decisions. Section 5 discusses the implications of the worker and firm decisions and provides two parametric examples of overeducation. Section 6 discusses about some implications of the model. Section 7 concludes.

II. Model Primitives

A. Worker, Education, and Human Capital

There is a continuum of workers who newly enter the labor market with different abilities (a). The ability a here should be interpreted as abstract ability or talent rather than specific abilities such as language or math ability. The ability of each worker is exogenous and permanent, so its distribution uniquely pins down the distribution of workers. We may further assume the worker ability a follows a distribution $F(\cdot)$, so $F(\cdot)$ also describes the distribution of workers.

The worker acquires human capital through education according to the human capital function, which is positive and depends only on education e.

$$h = h\left(e\right) \tag{1}$$

The interpretation of human capital here is productive knowledge or skills that can be learnt from education. Even without any formal education, workers possess certain low levels of knowledge and skills, so human capital is always positive. We assume knowledge acquisition and skill formulation through education are homogeneous across ability types, so that the amount of education determines the human capital level. We further assume the human capital function is increasing and weakly concave in the education level. The weakly concave human capital function rules out the possibility of accelerating human capital formation. Depending on the specific context, the rate of skill formation may be constant or diminishing. For simplicity, we also assume h is twice differentiable.

ASSUMPTION 1: $h'\left(e\right) > 0,\ h''\left(e\right) \leq 0,\ h\left(0\right) > 0$

While workers obtain human capital h through education, education also incurs cost c to the worker according to the cost function $c(\cdot)$, which depends on both ability a and education e. For simplicity, again we assume twice differentiability.

$$c = c(e, a) \tag{2}$$

The cost c here summarizes all the costs associated with education, including monetary cost and the psychic cost of efforts, as well as time. We follow the original Spence model and assume the cost of education is increasing in the amount of education e and decreasing in ability a. We further assume the cost of education is convex in e, and the cross derivative of e and a are negative. The rationale behind this assumption is that the cost of education grows faster as the education level gets higher (for instance, higher tuition and more efforts), while higher ability mitigates the accelerating cost.

ASSUMPTION 2: $\frac{\partial c}{\partial e}(a, e) > 0,\ \frac{\partial c}{\partial a}(a, e) < 0,\ \frac{\partial^2 c}{\partial a \partial e}(a, e) < 0$

The human capital formulation function $h\left(e\right)$ does not take ability a explicitly as an input. Such an assumption may seem unrealistic, since in practice one usually expects ability to play a positive role in human capital. However, we argue that this treatment is innocuous since the effect of ability is incorporated in the cost function of education. The positive effect of higher ability in human capital formation could be safely translated as lower cost of obtaining the same education. We closely follow the original Spence model, so adopt this version of modified assumption. Another reason for not including ability in the human capital function is more technical. As we will assume later, firms can observe the worker's human capital and education. For the signaling model to generate meaningful implications, the human capital function cannot include ability a, otherwise the firm can uniquely pin down the ability level from the worker's education and human capital.

B. Firm, Wage, Information, and Production

We assume there are two types of firms, low-paying firms (denoted by firm L) and high-paying firms (denoted by firm H). Following the specification of Kalemli-Ozcan, Ryder and Weil (2000), each type of firm pays a fixed wage per unit of human capital. In other words, L firms pay $w_L h$ to a worker with human capital h while H firms pay $w_H h$ to a worker with human capital h. H firms pay a higher wage than L firms, so $w_H > w_L$. This assumption of wage per unit of human capital is one of the major differences from the Spence model and allows for more flexibility in modeling the wage of a worker. Workers with more skills should be better compensated, so their wages increase with their skills (human capital h). Besides, on average all skills are valued with the same weight, so wages are paid on a per skill basis. Another thing to note is the fixed wage rates w_H and w_L. Fixed wage rates could arise in partial equilibria due to various reasons such as industry standard or simply time lag in adjustment. For now, we regard the wage rates as exogenous for the simplicity of analysis, so we are essentially investigating short-run partial equilibria.

A critical assumption is related to the information set that the firms receive about a worker. To illustrate the phenomenon of overeducation, we compare the information sets of firms in two scenarios. The first scenario has perfect information and no job market signaling. The information set of the firms in this scenario includes all the worker characteristics, (a, e, h). In contrast, imperfect information exists in the second scenario. Similar to the Spence model, we assume the firms observe only the education e of a worker when imperfect information persists.[3] Since we assume human capital h is uniquely determined by the level of education, effectively the firms observe both the education e and human capital h, but not the ability a.

We make a further simplifying assumption about the firms to liberate our analysis from excessive technical complexity. Since usually there are much more low-paying firms and relatively few high-paying firms, for simplicity we assume there

[3]Technically, firms also observe the worker's prior employment history, which is also an important signal of worker ability. However, since we focus on new labor market entrants with little prior work experience, this signal is rather weak.

are sufficiently many low-paying firms with varying production functions, so that any worker with any ability and education can find a job from L firms. In this way, we can treat L firms in a sense of outside option and focus on the decisions of H firms. Note that it does not mean all workers receive the same wage from L firms. Instead, their wages still depend on their skills. For simplicity, we use a representative firm H to represent the small number of H firms and consider the hiring decisions of firm H.

The production function of firm H for hiring a worker with ability a and human capital h is given below. Since human capital h is determined by education, the output is jointly determined by ability a and education e.

$$g(a, h) = g(a, h(e)) \tag{3}$$

Classical assumptions about the shape of the production function apply. The production function g is increasing in both ability a and human capital h; ability a and human capital h are complements in production; and there is non-accelerating marginal productivity in both inputs. Again, we allow for the possibility of constant returns to scale. Further we assume positive education is required for positive output, i.e., $g(a, h(0)) = 0$. This assumption reflects one source of minimum skill requirement by the high-paying firm H: the technology of the firm is sophisticated and requires at least some formal education for the worker to operate. For simplicity, we assume the production function is twice continuously differentiable.

ASSUMPTION 3: $\frac{\partial g}{\partial a} > 0,\ \frac{\partial g}{\partial h} > 0,\ \frac{\partial^2 g}{\partial a \partial h} > 0,\ \frac{\partial^2 g}{\partial a^2} \leq 0,\ \frac{\partial^2 g}{\partial h^2} \leq 0,\ g(a, h(0)) = 0$

III. Worker Optimization and Choice

In this section, we discuss the worker's objectives and general choices of education without considering the firm's information set. The workers make choices about job types and education levels to achieve a net payoff as high as possible.

A. Worker's Payoff

The workers are risk-neutral and care about their net payoff. Since there are two types of jobs, the worker has two versions of payoffs:

$$u_\theta(e, a) = w_\theta h(e) - c(e, a),\ \theta \in \{\ L, H\} \tag{4}$$

The worker's objective is to choose an education level and job type to maximize her net payoff. For our analysis, we propose a regularity assumption about the relationship between $h(e)$ and $c(e, a)$:

ASSUMPTION 4:

$$\frac{\partial u_\theta}{\partial e}(0,a) = w_\theta \frac{\partial h}{\partial e}(0,a) - \frac{\partial c}{\partial e}(0,a) > 0$$
$$\lim_{e\to\infty} \frac{\partial u_\theta}{\partial e}(e,a) = \lim_{e\to\infty} [w_\theta \frac{\partial h}{\partial e}(e,a) - \frac{\partial c}{\partial e}(e,a)] < 0$$

This assumption means when the education level is very low, any worker will obtain a higher net payoff in both L jobs and H jobs if she increases her education, but the reverse when the education level is already high. Since the human capital function is concave and the cost of education function is convex, it is innocuous to impose this restriction. Also note that if we instead assume the reverse, then the net payoff will decrease with education at any education level for some worker, so they will choose zero education. This result echoes with the Spence model and explains undereducation of some workers. It is an interesting result, but it is not our focus in this paper.

Under this regularity assumption, there is an optimal amount of education for each type of job that that generates the highest net payoff for a worker with ability a.

PROPOSITION 1: *For a worker with ability a, there are education levels $e_L^*(a)$ and $e_H^*(a)$ that solve the following maximization problems respectively. Additionally, $e_H^*(a) \geq e_L^*(a)$ and both increase with a.*

$$e_L^*(a) = \underset{e}{\operatorname{argmax}}\ u_L(e,a) = w_L h(e) - c(e,a)$$
$$e_H^*(a) = \underset{e}{\operatorname{argmax}}\ u_H(e,a) = w_H h(e) - c(e,a)$$

Furthermore, if possible, all workers have incentives to achieve their $max\ u_H$, since it is at least as large as $max\ u_L$.

PROPOSITION 2: *For a worker with ability a,*

$$max\ u_H(e,a) = u_H(e_H^*(a), a) \geq max\ u_L(e,a) = u_L(e_L^*(a), a)$$

Additionally, $\Delta(a) = max\ u_H(e,a) - max\ u_L(e,a)$ increases with ability a.

As discussed above, a worker can always find a job among L firms, so a worker can always achieve her $max\ u_L = u_L(e_L^*(a), a)$. However, as we will see later, a worker may be ineligible for an H job. Let $\delta_{e,a}$ denote whether a worker with ability a and education e is eligible for an H job. $\delta_{e,a} = 1$ if the worker is eligible, 0 otherwise. Then we can formally write down the net payoff to a worker with ability a who chooses education e:

$$U(e,a) = max\,(max\ u_L(e,a), \delta_{e,a} u_H(e,a))$$

At any education level, the worker checks if she is eligible for job H and her net payoff from job H if eligible. Then she compares this net payoff with her outside option and chooses the higher one.

B. Worker's Breakeven Education

A worker's net payoff from an H job first increases with her education, then decreases after the maximum point e_H^*. Since $max\ u_H(e,a) \geq max\ u_L(e,a)$, technically there is a range of education levels that give higher net payoffs to the worker. Additionally, there is a breakeven education level higher than e_H^* that gives the worker the same net payoff between an H job and the outside option. This breakeven education level is the maximum that a worker is willing to acquire for an H job. If the education requirement for an H job exceeds this breakeven education, then an H job will be less desirable than the best L job. Proposition 3 states the existence of this breakeven education level and its properties.

PROPOSITION 3: *For a worker with ability a, there are education levels $\underline{e}_b(a) \leq e_H^*(a) \leq \overline{e}_b(a)$ such that $u_H(\overline{e}_b(a)) = max\ u_L(e,a)$ and $u_H(e,a) < max\ u_L(e,a)$ for any $e < \underline{e}_b(a)$ or $e > \overline{e}_b(a)$. $\underline{e}_b(a)$ is decreasing with a, while $\overline{e}_b(a)$ is increasing with a. Note that $\underline{e}_b(a)$ may be zero, and in this case $u_H(\underline{e}_b(a))$ may not be equal to $max\ u_L(e,a)$.*

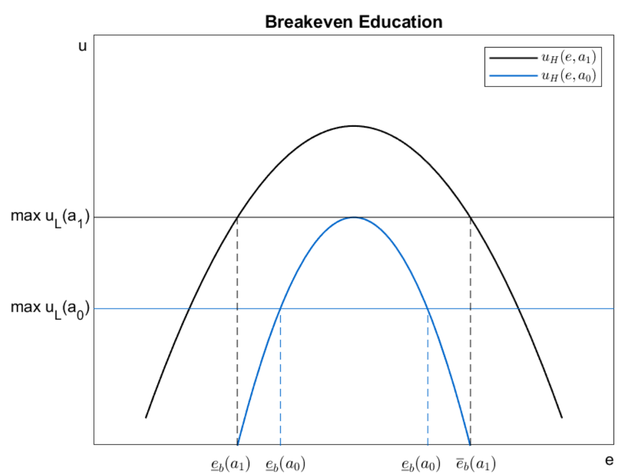

FIGURE 1. ILLUSTRATION OF BREAKEVEN EDUCATION

This proposition is a critical result for our analysis. Figure 1 is a visual illustration of its implications. It establishes that there is a range of education levels $(\underline{e}_b(a), \ \overline{e}_b(a))$ that each worker is willing to acquire to obtain an H job. When the education level is low, the worker can only receive a low wage from an H job, making the net payoff from an H job lower than what she can always receive from the outside option of L jobs. When the education level increases to a certain range, the wage from an H job is high enough so that the worker can receive a higher net payoff from H jobs and reach the maximum at e_H^*. However, as the education level increases beyond the breakeven level, the cost of education dominates, and the outside option again becomes better. This analysis has two important implications. First, the worker wants to choose an education level as close to e_H^* as possible. Second, even if it takes more education than the worker's optimal level e_H^* to get an H job, she is also willing to do so as long as it does not exceed her breakeven education e_b. We will see later that the second implication is the source of potential overeducation in the signaling equilibrium.

IV. Objective and Education Requirement of Firm H

The other side of the job market is the firms. With the simplifying assumption about the firms, we can focus on firm H and its decisions. This section provides some general analysis of the decisions of firm H without considering the information set. The results derived in this section will serve as critical preliminaries for the comparison between perfect information and signaling equilibria.

The objective of firm H is to maximize its total profit. The profit from hiring a worker with ability a and education e is given by:

$$V(a,e) = g(a, h(e)) - w_H h(e) \tag{5}$$

The worker's ability a affects firm H's profit only through the production function and higher ability increases profit. In contrast, education e affects profit through the production function and the cost of human capital in opposite directions. A better educated worker produces more output but is also more expensive.

To maximize total profit, firm H will hire a worker with ability a and education e as long as the worker can bring positive profit, i.e., $V(a,e) \geq 0$. Therefore, for each ability level, the firm has an education requirement.

PROPOSITION 4: *At each ability level a, if firm H can observe all worker characteristics, it will have a range of education requirement* $(\underline{e}_r(a), \overline{e}_r(a))$. *This range could be empty and becomes wider as a increases.*

This proposition has some critical implications. The first implication is the constraint on the worker's job choice. If firms can observe worker ability a (i.e., under perfect information), then firm H will reject workers with education levels off the required education range $\underline{e}_r(a)$. Although each worker wants to obtain her e^*_H education and an H job, the minimum education requirement $\underline{e}_r(a)$ by firm H may be higher than the breakeven education $\overline{e}_b(a)$. It is even possible that firm H does not hire workers with certain low ability levels at all. In either case, some workers with low ability levels are rejected by firm H under perfect information. In contrast, under imperfect information firms never observe worker ability a, but infer about it from education signals. Firm H will accept a worker only if the worker's education satisfies the education requirement for her inferred ability. In this scenario, the low-ability workers have incentives to use education to falsify their ability in signaling equilibria. Another implication is from the width of the education requirement. While firm H has a minimum education requirement, it also has a maximum education limit.[4] As ability level increases, firm H accepts a wider range of education e in the perfect information equilibrium. Workers with higher ability are subject to looser constraints by firm H and enjoy more flexibility in their education choices. In the best case, workers with very high abilities may choose their optimal education e^*_H.

V. Overeducation in Signaling Equilibria

A. *Possibility of overeducation*

The previous sections developed the worker's and firm H's objective and decisions. However, the interactions between workers and firms, either in a world with perfect information or in a signaling world, are minimally described. It is difficult to draw any additional meaningful insights into the interaction between workers and firms in this most general form. The barrier to further analysis is the obscure relationship between the worker's breakeven education $(\underline{e}_b, \overline{e}_b)$ and firm H's education requirement $(\underline{e}_r, \overline{e}_r)$. Without further specification of the functions and wages in the model, the relationship between these two sets of quantities is obscure. This flexibility of our analysis allows for various equilibria under both perfect information and job market signaling.

One possibility is the overeducation of middle-ability workers in the signaling equilibrium compared to the perfect information world. This possibility occurs under certain specifications of the functional forms and fixed wage rates, as described by Figure 2. In this scenario, the required education for a_0 is $\underline{e}_r(a_0) > \overline{e}_b(a_0)$, while the optimal education for a_1 is $e^*_H(a_1) > \underline{e}_r(a_1)$.

Under perfect information, firm H can observe the ability of each worker and has different education requirements for different abilities. a_1 workers can choose H jobs and her optimal education. a_0 workers also want to get H jobs with $e^*_H(a_0)$ education, but the education requirement $(\underline{e}_r, \overline{e}_r)$ for the low ability workers a_0

[4] This upper bound is really just a technical result for the sake of completeness. In practice, as we can see in the examples below, the upper limit in education requirement rarely affects worker and firm decisions.

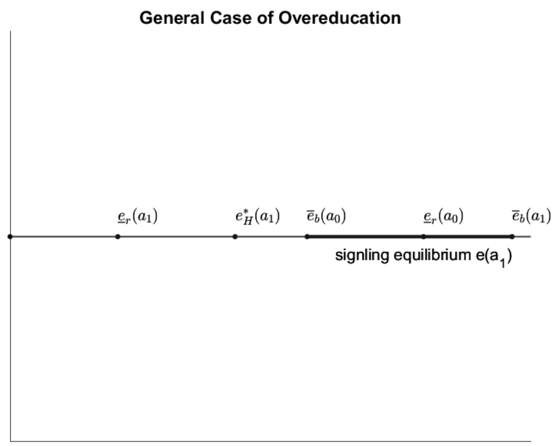

FIGURE 2. POSSIBILITY OF OVEREDUCATION

exceeds their breakeven education level $(\underline{e}_b, \overline{e}_b)$. It is even possible that firm H rejects all workers below some low ability level, as hiring them cannot generate positive profits. Those workers must stick to L jobs. In contrast, under imperfect information, firm H can only observe education and workers need to use education as a signal for their ability. In this case, some of the workers in L jobs with ability a_0 have incentives to pretend to have ability by choosing $e^*_H(a_1) < \overline{e}_b(a_0)$, the optimal education level of workers with higher ability $a_1 > a_0$. Since the firm cannot distinguish a_1 workers from a_0 workers at $e^*_H(a_1)$ education, it is reluctant to hiring workers with this education. To signal their higher ability and receive the high net payoff from H jobs, a_1 workers will now choose education $e(a_1)$ higher than the a_0 breakeven education $\overline{e}_b(a_0)$ but lower than their own breakeven education $\overline{e}_b(a_1)$, i.e., $e^*_H(a_1) < \overline{e}_b(a_0) < e(a_1) < \overline{e}_b(a_1)$.

This simple exercise has critical implications: middle ability workers may be harmed in the signaling equilibrium. Compared to the perfect information equilibrium, a_1 workers are overeducated and earn lower net payoff. In contrast, since the optimal education $e^*_H(a)$ increases with ability a, workers with very high abilities will have $e^*_H(a) > \overline{e}_b(a_0)$. Therefore, a_0 workers will not try to pretend to possess high ability, so workers with very high abilities are not affected by the signaling behavior of a_0 workers. Finally, workers with very low abilities still stick to their L jobs and will not try to send signals. The upper breakeven educations $\overline{e}_b(a)$ of such workers are lower than the lower breakeven education $\underline{e}_b(a)$ of any worker at H jobs, so they cannot improve their net payoff by pretending to possess higher abilities.

On the other hand, it is also possible to have a signaling equilibrium without overeducation. As firm H becomes more productive, it's easier for the firm to earn positive profit, so the range of education requirements $(\underline{e}_r(a), \overline{e}_r(a))$ becomes wider for each ability level. Therefore, it is possible for the range of education requirements to include the optimal education level $e_H^*(a)$ for all workers. In this case, all workers will obtain H jobs and receive their optimal education $e_H^*(a)$, both in the perfect information equilibrium and the signaling equilibrium. The unobservability of worker ability does not create overeducation in the signaling equilibrium, neither does it harm any worker.

In the following parts of this section, we provide two numerical examples to illustrate this possibility of overeducation in the signaling equilibrium.

B. Signaling Equilibrium with Overeducation-Numeric Example

For illustrative purposes and simplicity of analysis, we assume the following functional forms and fixed wage rates. In the perfect information equilibrium, firm H can observe all the worker characteristics. It is a good starting point for studying the signaling equilibrium.

Worker ability: $a \in [\underline{a}, \overline{a}] = [1, 10]$
Human capital function: $h(e) = 1 + e$
Cost of education: $c(e, a) = \frac{1}{a}e^2$
Production function: $g(a, h) = g(a, h(e)) = a(h-1) = ae$
Wage rates: $w_H = 2,\ w_L = 1$

Therefore, the net payoff functions to the workers are:

$$
\begin{aligned}
u_H(e, a) &= w_H h(e) - c(e, a) = 2e + 2 - \frac{1}{a}e^2 \\
u_L(e, a) &= w_L h(e) - c(e, a) = e + 1 - \frac{1}{a}e^2
\end{aligned} \tag{6}
$$

The profit of firm H hiring a (e, a) worker is:

$$
V(e, a) = g(a, h(e)) - w_H h(e) = ae - 2e - 2 \tag{7}
$$

We first consider the worker's maximization problem for L jobs and H jobs. The FOCs are:

$$
\begin{aligned}
\frac{\partial u_H}{\partial e}(e, a) &= 2 - \frac{2}{a}e = 0 \\
\frac{\partial u_L}{\partial e} &= 1 - \frac{2}{a}e = 0
\end{aligned} \tag{8}
$$

These FOCs give the optimal education level for L jobs and H jobs:

(9) $$e_H^*(a) = a, \quad e_L^*(a) = \frac{a}{2}$$

The worker can always obtain their highest net payoff at L jobs, which is:

(10) $$max\ u_L(a) = u_L(e_L^*(a), a) = e_L^*(a) + 1 - \frac{1}{a} e_L^*(a)^2 = \frac{a}{4} + 1$$

At each education level, the worker compares her $u_H(e,a)$ with $max u_L(a)$, and the range of education for the worker to desire H jobs is given by the condition:

(11) $$u_H(e,a) \geq max\ u_L(a) \Leftrightarrow 2e + 2 - \frac{1}{a}e^2 \geq \frac{a}{4} + 1$$

Solving the inequality gives $\underline{e}_b(a) = a - \frac{a}{2}\sqrt{\frac{4}{a}+3}$ and $\overline{e}_b(a) = a + \frac{a}{2}\sqrt{\frac{4}{a}+3}$. The worker desires an H job only within this range of education. Note that this range includes $e_H^*(a) = a$.

After we are clear about the worker decisions, we consider the education requirements by firm H. Firm H requires the profit from hiring a worker is non-negative, so the condition is given by:

(12) $$V(e,a) = g(a, h(e)) - w_H h(e) = ae - 2e - 2 \geq 0$$

This condition gives the inequality $(a-2)e \geq 2$. This inequality has interesting implications. Under perfect information, firm H will not hire workers with ability $1 \leq a \leq 2$, regardless of their education. For workers with ability $a > 2$, firm H has a minimum education requirement $\underline{e}_b(a) = \frac{2}{a-2}$. In this special case, the maximum education $\overline{e}_b(a) = \infty$ for $a > 2$.

The next step is to solve for the worker's education and jobs in the perfect information equilibrium. Based on previous discussions, four decision rules determine the perfect information equilibrium. Solutions to the respective conditions give the cut-off ability levels. The perfect information equilibrium is illustrated in Figure 3.

1) Workers with $1 \leq a < 2$ cannot get an H job and receive $e_L^*(a)$ education and their best L jobs.
2) Workers with $\underline{e}_r(a) > \overline{e}_b(a)$ also choose to receive $e_L^*(a)$ education and their best L jobs, since the education requirement of H jobs is higher than their breakeven education. The range of ability is $2 \leq a < 2.4$. [5]

[5]approximated by numerical methods, since the analytical form is complicated to solve.

3) Workers with $e_H^*(a) < \underline{e}_r(a) \leq \overline{e}_b(a)$ choose H jobs. However, they cannot choose their optimal education $e_H^*(a)$, so they choose the closest possible education $\underline{e}_r(a)$. The range of ability is $2.4 \leq a < 1 + \sqrt{3}$.

4) Workers with $e_H^*(a) \geq \underline{e}_r(a)$ choose H jobs and their optimal education $e_H^*(a)$. The range of ability is $1 + \sqrt{3} \leq a \leq 10$.

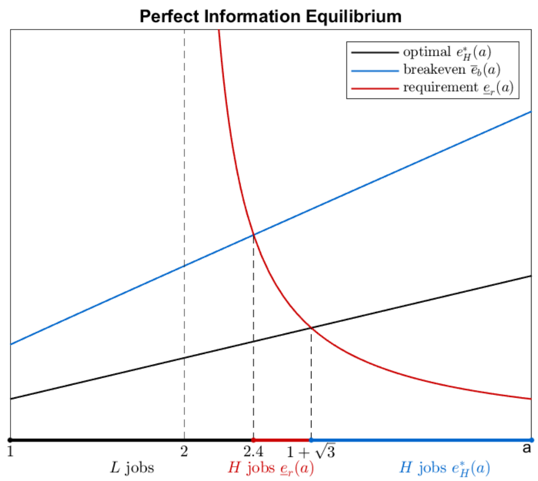

Figure 3. Perfect Information Equilibrium

Next, we study the signaling equilibrium with the same set of functions and wage rates. In the signaling equilibrium, firm H can only observe the education of workers. As the ability level is private information, some workers at L jobs have incentives to pretend to possess higher ability by choosing the education choice of workers with higher ability. We refer to such workers as imposters and their signaling education as imposter education in later contexts. The imposters will not choose arbitrarily high imposter educations, but follow the rules below:

1) They will only pretend to be workers that can and will choose H jobs. This implies that they will pretend to possess abilities at least $a = 2.4$.

2) They will not exceed their own breakeven education $\overline{e}_b$, otherwise getting an H job would give a lower net payoff than their L job.

The first rule gives the "pretense function", the relationship between the pretense ability level and the education level the imposter should choose. It is also the education level chosen by workers really with the corresponding ability. As derived in the previous section and shown in Figure 4, the education choice of a worker at job H is:

$$e(a) = \begin{cases} \underline{e}_r(a) = \frac{2}{a-2} \text{ if } 2.4 \le a < \sqrt{3}+1 \\ e(a) = e_H^*(a) = a \text{ if } a \ge \sqrt{3}+1 \end{cases} \tag{13}$$

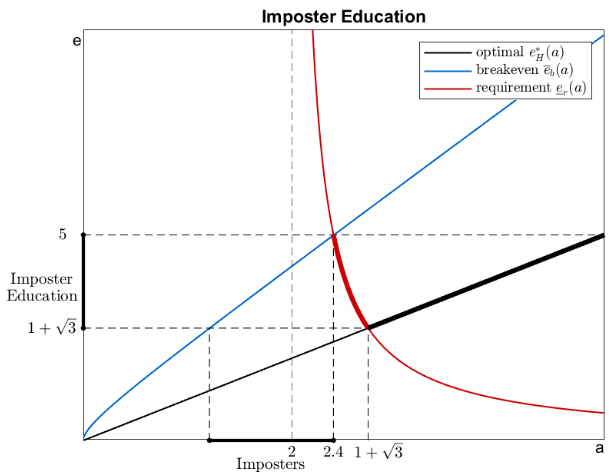

FIGURE 4. EDUCATION AT H JOBS WITH IMPOSTERS

This pretense function has an important implication about the ability range of imposters. Figure 4 shows that lowest education choice occurs at the ability level $a = \sqrt{3}+1$ with $e(a) = \sqrt{3}+1$. It is the minimum education that an imposter should receive to pretend to have an ability level eligible for job H. Since an imposter will never receive education higher than her own $\overline{e}_b$, this minimum education gives the lower bound $\underline{a}_{im}$ of the ability of an imposter. Solving for the boundary condition, we can derive the value of $\underline{a}_{im}$:

$$\overline{e}_b(\underline{a}_{im}) = \overline{e}_b(a) = a + \frac{a}{2}\sqrt{\frac{4}{a}+3} = \sqrt{3}+1 \Rightarrow \underline{a}_{im} = 1.211^{6} \tag{14}$$

Any worker with ability lower than $\underline{a}_{im}$ will not be an imposter since the pretense education is higher than their breakeven education. Additionally, we can also derive the upper bound $\overline{a}_{im}$ of the ability of imposters. The upper bound is simply the cut-off ability for H jobs and L jobs. Therefore, the ability range of imposters is $(\underline{a}_{im}, \overline{a}_{im}) = (1.211, 2.4)$.

[6] It's also solved using numerical methods. Although the analytical solution exists and can be computed, which is $6+4\sqrt{3}-2\sqrt{17+10\sqrt{3}}$, the approximation is more meaningful for interpretation.

Another critical result from the pretense function is the range of education levels that imposters may choose. The maximum pretense education that an imposter may choose is her $\overline{e}_b$, so the ability range of imposters $(\underline{a}_{im}, \overline{a}_{im})$ also generates the imposter education range $(\underline{e}_{im}, \overline{e}_{im}) = (\overline{e}_b(\underline{a}_{im}), \overline{e}_b(\overline{a}_{im}))$, which is $(1+\sqrt{3}, 5)$ in this case.

Although imposters have incentives to pretend to have high ability, firm H and high-ability workers will counter-act in the signaling equilibrium. On the one hand, firm H will refrain from hiring workers with the imposter education range, since it cannot distinguish imposters from non-imposters and hiring imposters generates negative profits. On the other hand, to signal their identity as non-imposters, non-imposters will choose education levels higher than the imposter education range. Figure 5 shows a visual illustration of the signaling equilibrium. Imposters will stay at L jobs and choose their $e_L^*(a)$ in the signaling equilibrium. Non-imposters whose original education levels fall into the imposter education range will now choose the upper bound of the imposter education range. Meanwhile, even higher-ability workers (in this example with ability $a > 5$) will stick to their optimal education since they are not affected by imposters.

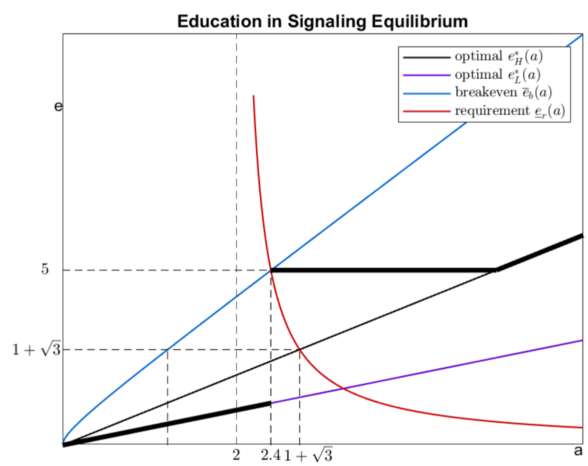

Figure 5. Education in the Signaling Equilibrium

The most important observation here is that although non-imposters can effectively distinguish themselves from imposters by receiving more education, the middle workers (within the imposter education range) are worse off in the signaling equilibrium.

C. Signaling Equilibrium without Overeducation

While the previous example illustrated the possibility of middle ability workers being harmed by overeducation in the signaling equilibrium, this section provides a contrasting example. The example in this section represents the case when the lack of information about the worker ability a does not cause overeducation in the signaling equilibrium. We assume the following functional forms for this example. Note that the only distinction from the previous example is the production function.

Worker ability: $a \in [\underline{a}, \overline{a}] = [1, 10]$
Human capital function: $h(e) = 1 + e$
Cost of education: $c(e, a) = \dfrac{1}{a} e^2$
Production function: $g(a, h) = g(a, h(e)) = 4a^{1/2}(h-1)^{1/2} = 4a^{1/2}e^{1/2}$
Wage rates: $w_H = 2,\ w_L = 1$

Since the worker functions are the same as in the previous example, the breakeven educations are the same. $(\underline{e}_b(a), \overline{e}_b(a)) = \left(a - \frac{a}{2}\sqrt{\frac{4}{a}+3}, a + \frac{a}{2}\sqrt{\frac{4}{a}+3}\right)$. However, firm H's education requirement is different in this example, given by:

$$V(e, a) = g(a, h(e)) - w_H h(e) = 4a^{1/2}e^{1/2} - 2e - 2 \geq 0 \tag{15}$$

The condition gives the inequality $e^2 + (2 - 4a)e + 1 \leq 0$. This inequality determines the education requirement by firm H if it can observe worker ability. When $a^2 - a < 0$, which means $0 < a < 1$, $e^2 + (2 - 4a)e + 1 \leq 0$ never holds, so firm H does not hire any worker with ability range $a \in (0, 1)$. On the other hand, when $a \geq 1$, the range of education requirement is given below. Firm H will only hire workers with education levels in this range.

$$(\underline{e}_r(a), \overline{e}_r(a)) = \left(2a - 1 - 2\sqrt{a^2 - a}, 2a - 1 + 2\sqrt{a^2 - a}\right) \tag{16}$$

This result contrasts with the education requirement in the previous example. Due to the concavity of the production function, firm H now not only has a minimum education requirement $\underline{e}_r(a)$, but also has a maximum education requirement $\overline{e}_r(a)$.[7] Workers with too low education have low productivity, but workers with too high education are too costly. This maximum education requirement implies the possible existence of labor market penalty for overeducation, which is consistent with empirical findings.

Based on the above discussions, the decision rule in the perfect information

[7] However, in our example here, it does not affect worker and firm decisions, as we see below.

equilibrium is surprisingly simple and stated below. Figure 6 is a visual representation of the perfect information equilibrium.

1) When $a \geq 1$, $\underline{e}_r(a) \leq e_H^*(a) \leq \overline{e}_r(a)$. All workers can get H jobs and receive their optimal education $e_H^*(a)$.

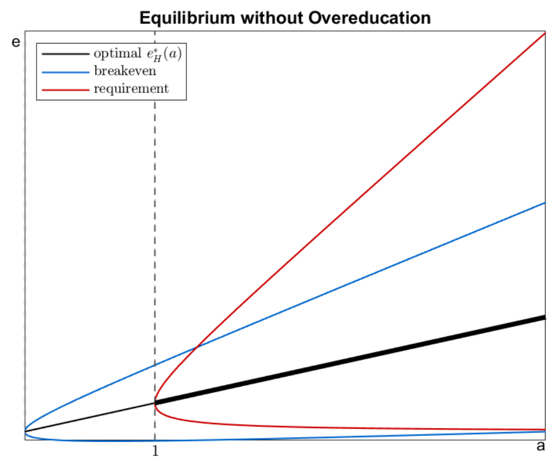

Figure 6. Equlibrium without Overeducation

We then consider the signaling equilibrium when firm H cannot observe the worker ability. A crucial observation from the perfect information equilibrium is that all workers receive their best outcome, so nobody has incentives to pretend to have other people's abilities. Therefore, no worker will try to send signals through overeducation or undereducation and everyone still receives their optimal education e_H^*. The signaling equilibrium is exactly the perfect information equilibrium. The lack of information about the worker ability does not distort the labor market.

VI. Discussions

A. Implications for Empirical Analysis

The analysis in this section has several implications for the empirical estimation of the return to education. The standard practice focuses on wages and estimates the famous Mincer (1958) regression. Unlike our context of the starting job, employers gradually learn the worker ability as workers gain more work experience (Kahn and Lange, 2014), and the observed wage data used in the Mincer regression usually also depends on worker ability. Our analysis suggests that workers self-select into different education with a positive correlation, so the standard empirical practice tends to overestimate the wage return to education if it does not address the self-selected education by ability. The analysis about the signaling equilibrium suggests that overestimation may also occur in empirical studies that try to address the varying ability by investigating the locale of anchor education levels and thus computing the local average treatment effect. The example shown in section 5.2 implies that a large portion of workers with medium education are not as talented as expected, so overestimation occurs around this medium education level.

The most important implication of our analysis is that overeducation only occurs within some middle-range education, so the lower and higher tails of the worker education distribution are not affected. Therefore, compared to the standard Mincer mean regression, a quantile regression will better address the bias issue. Even with mean regressions, our analysis also shows that it may not always overestimate the return to education. If the distortion of education from signaling is not strong, or the empirical analysis as of the case presented in section 5.3, then overestimation will be less of a concern. Alternatively, if the range of education attainment is wide enough and we can make parametric assumptions on the functional form of the return to education, then we can still address the overestimation bias with a mean regression on the tails (This is essentially an identification in the infinity approach).

B. *Signaling Power of Different Educations*

The essential source of overeducation in the signaling equilibrium, as shown in our analysis, is the imperfect information about worker abilities and the weak signaling power of education. Workers have full knowledge of their own abilities, while firms can only infer about it from education signals. Meanwhile, the education per se in our context does not convey any direct information about worker ability. It does not provide any assessment of worker ability, and serves as a signal for ability only through the worker's education cost function and payoff maximization. Because of the imperfect information and the weak signaling power of education, it is possible for some workers to pretend to have high abilities by over-investing in education.

In our analysis, we have only one type of education with weak signaling power. This analysis could be extended to other types of education with medium or strong signaling power. Under imperfect information, if certain types of education are good proxies of ability (For instance, Arcidiacono, Bayer and Hizmo (2010) show that college graduation directly reveals worker ability in the labor market in the United States.), then the imperfect information will be alleviated and the signaling equilibrium will be closer to that in the perfect information context with less overeducation. The difference in signaling power could explain the variation in the degrees of overeducation across education types, employment sectors, and education systems. It also provides a policy suggestion to address the overeducation problem.

VII. Conclusions

This paper focuses on partial equilibria with fixed wages and is a generalization of the Spence (1973) signaling model. Instead of treating education as a pure undesirable good with zero productivity improvement, our model incorporates the human capital effect of education. This human capital effect is a second channel through which education affects job market outcomes. The trade-off between wage premiums and the cost of education entails decisions for maximization problems. We offer a set of general properties of the decisions of workers and firms. Based on those properties, both the perfect information and signaling equilibria are jointly determined by the interactions between worker and firm choices. In the most general form, since the production function and the education cost functions are not intrinsically tied, the comparison between the signaling equilibria and the perfect information equilibria is not fully clear. Depending on the exact production and education cost functions, various equilibria could arise.

The framework in our model allows for the flexibility of fitting into different contexts. On the one hand, under some reasonable functional forms, overeducation still occurs in the signaling equilibrium. In this case, the most important result is that overeducation in the signaling equilibrium only occurs for middle workers as they choose excessive education to signal their ability. Another observation is that, in contrast to the Spence model, all workers in this model choose some education. This observation is a direct corollary of the introduction of human capital, as education is no longer a pure cost. On the other hand, under other conditions, the unobservability of the worker ability does not distort the job market, and the signaling equilibrium is exactly the perfect information equilibrium. This result contrasts with the Spence model and suggests that the inefficiency in the job market caused by the imperfect information may be alleviated by the human capital channel.

One important empirical implication of our analysis is that the standard Mincer regression tends to overestimate the return to education. Workers self-select into difference education by their ability, and this positive correlation between education and ability causes overestimation. Furthermore, overestimation may also occur in estimates of the local average treatment effect due to the overeducation of middle workers in the signaling equilibrium. The overestimation bias is stronger when overeducation is more severe. However, our analysis also suggests that overeducation usually only occurs in a middle range, so a quantile regression can better address this problem. Alternatively, if the distribution of education attainment is wide enough, then mean regressions on the tails could also address the overestimation problem if parametric assumptions have been made about the return to education.

The essential source of overeducation, as our analysis shows, is the imperfect information about worker abilities. In our setting, firms cannot observe worker abilities. Similarly, education per se does not provide any information about worker ability. Because of this imperfect information and weak signaling power of education, it is possible for some workers to pretend to have higher abilities by over-investing in education. If firms can directly observe worker ability or some types of education provide good assessment of worker ability, then sending signals by overeducation is impossible, and overeducation in the signaling equilibrium will disappear. This discussion suggests that better designs of the education system with comprehensive and transparent ability assessment can alleviate overeducation due to signaling.

This paper is subject to several limitations that may be addressed in future works. Our analysis takes wage rates as given and fixed, which is essentially a partial equilibrium approach. Hence, the implications from our model should be interpreted as short-run equilibria. This problem may be addressed in a general equilibrium analysis that also models the determinants of wages. Additionally, we substantially simplify the analysis by dividing firms into high-pay firms and low-pay firms and essentially treat low-pay firms as the outside option. Alternatively, more heterogeneity across firms could be introduced by modeling the distribution of different firms. This framework has the potential to be developed into a more sophisticated structural analysis by combining the extensions with the estimated production and cost functions in specific settings. This direction may be addressed in future works.

REFERENCES

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics*, 2(4): 76–104.

Baert, Stijn, Bart Cockx, and Dieter Verhaest. 2013. "Overeducation at the start of the career: Stepping stone or trap?" *Labour Economics*, 25: 123–140.

Barrera-Osorio, Felipe, and Hernando Bayona-Rodríguez. 2019. "Signaling or better human capital: Evidence from Colombia." *Economics of Education Review*, 70: 20–34.

Becker, Gary S. 1964. *Human capital: a theoretical and empirical analysis, with special reference to education.* Vol. no. 80.;no. 80;, New York:National Bureau of Economic Research; distributed by Columbia University Press.

Bleemer, Zachary, Meta Brown, Donghoon Lee, Katherine Strair, and Wilbert van der Klaauw. 2021. "Echoes of rising tuition in students' borrowing, educational attainment, and homeownership in post-recession America." *Journal of Urban Economics*, 122: 103298.

Castro, Juan, Lorena Ortega, Gustavo Yamada, and David Mata. 2022. "The Magnitude and Predictors of Overeducation and Overskilling in Latin America: Evidence from PIAAC." *SSRN Electronic Journal*.

Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss. 2009. "Investment in Schooling and the Marriage Market." *American Economic Review*, 99(5): 1689–1713.

Cultrera, L., B. Mahy, F. Rycx, and G. Vermeylen. 2022. "Educational and skills mismatches: unravelling their effects on wages across Europe." *Education Economics*, 30(6): 561–573.

European Commission. 2017. *Employment and social developments in Europe: annual review 2017.* Luxembourg:Publications Office of the European Union.

Garcia-Mainar, Inmaculada, and Victor M. Montuenga. 2019. "The signalling role of over-education and qualifications mismatch." *Journal of Policy Modeling*, 41(1): 99–119.

Groot, Wim, and Henriëtte Maassen van den Brink. 2000. "Overeducation in the labor market: a meta-analysis." *Economics of Education Review*, 19(2): 149–158.

Kahn, Lisa B., and Fabian Lange. 2014. "Employer learning, productivity, and the earnings distribution: Evidence from performance measures." *The Review of economic studies*, 81(4): 1575–1613.

Kalemli-Ozcan, Sebnem, Harl E Ryder, and David N Weil. 2000. "Mortality decline, human capital investment, and economic growth."

Marioni, Larissa da Silva. 2021. "Overeducation in the labour market: evidence from Brazil." *Education Economics*, 29(1): 53–72.

McGuinness, Seamus. 2006. "Overeducation in the Labour Market." *Journal of Economic Surveys*, 20(3): 387–418.

Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution." *The Journal of political economy*, 66(4): 281–302.

Naven, Matthew, and Daniel Whalen. 2022. "The signaling value of university rankings: Evidence from top 14 law schools." *Economics of Education Review*, 89: 102282.

Spence, Michael. 1973. "Job Market Signaling." *The Quarterly Journal of Economics*, 87(3): 355.

Tsang, Mun C., Russell W. Rumberger, and Henry M. Levin. 1991. "The Impact of Surplus Schooling on Worker Productivity." *Industrial relations (Berkeley)*, 30(2): 209–228.

Mathematical Appendix

A1. Glossary

Symbol	Definition
a	worker ability/talent
e	worker education
h	human capital: productive knowledge and skills
$h(e)$	human capital accumulation function
c	cost of education
$c(e,a)$	education cost function
$\theta \in \{H, L\}$	firm type: high-paying & low-paying
w_H	wage in H firms per human capital
w_L	wage in L firms per human capital
$g(a,h)$	firm H's production function
$u_\theta(e,a)$	worker's hypothetical net payoff in θ firms
e_L^*	worker's optimal education for L jobs
e_H^*	worker's optimal education for H jobs
$\delta_{e,a}$	indicator of an (e,a) worker's eligibility for an H job
$U(e,a)$	actual net payoff of an (a,e) worker t
$(\underline{e}_b(a), \overline{e}_b(a))$	worker's breakeven education between H and L jobs
$V(a,e)$	profit to the H firm from hiring an (a,e) worker
x	fixed cost of education
$(\underline{e}_r(a), \overline{e}_r(a))$	min and max education requirement by the H firm

A2. Proof of Proposition 1

Consider the worker's net payoff function at H and L jobs:

$$u_\theta(e,a) = w_\theta h(e,a) - c(e,a)\ ,\theta \in \{L, H\}$$

We want to know how $u_\theta(e,a)$ changes with e and a by taking the first-order derivatives:

$$\text{(A1)} \qquad \frac{\partial u_\theta}{\partial a}(e,a) = w_\theta \frac{\partial h}{\partial a}(e,a) - \frac{\partial c}{\partial a}(e,a) > 0$$

Equation (A1) is positive since $\frac{\partial h}{\partial a}(e,a) > 0$ and $\frac{\partial c}{\partial a}(e,a) < 0$. The positive first-order derivative suggests the worker's net payoff increases with ability level a.

$$\text{(A2)} \qquad \frac{\partial u_\theta}{\partial e}(e,a) = w_\theta \frac{\partial h}{\partial e}(e,a) - \frac{\partial c}{\partial e}(e,a)$$

The sign of equation (A2) is not immediately determined, since both $\frac{\partial h}{\partial e}(e,a) > 0$ and $\frac{\partial c}{\partial e}(e,a) > 0$. To investigate how it changes with the education level e, we can take the second-order derivative:

$$\text{(A3)} \qquad \frac{\partial^2 u_\theta}{\partial e^2}(e,a) = w_\theta \frac{\partial^2 h}{\partial e^2}(e,a) - \frac{\partial^2 c}{\partial e^2}(e,a) < 0$$

Equation (A3) is negative because $\frac{\partial^2 h}{\partial e^2}(e,a) \leq 0$ and $\frac{\partial^2 c}{\partial e^2}(e,a) > 0$. It suggests that the first-order derivative $\frac{\partial u_\theta}{\partial e}(e,a)$ decreases with education e. Additionally, by regularity Assumption 4:

$$\text{(A4)} \qquad \frac{\partial u_\theta}{\partial e}(0,a) > 0,\ \lim_{e\to\infty} \frac{\partial u_\theta}{\partial e}(e,a) < 0$$

Therefore, there exists e^*_θ such that $\frac{\partial u_\theta}{\partial e}(e^*_\theta, a) = 0$. The optimal education levels from the worker's perspective are e^*_H for H jobs and e^*_L for L jobs.

Next, we study how e^H_* compares to e^L_* and how they change with ability a. It is easy to see that $e^H_* > e^L_*$. By the FOCs:

$$\text{(A5)} \qquad \frac{\partial u_L}{\partial e}(e^*_L, a) = w_L \frac{\partial h}{\partial e}(e^*_L, a) - \frac{\partial c}{\partial e}(e^*_L, a) = 0$$

Since $w_H > w_L$ and $\frac{\partial h}{\partial e}(e^*_L, a) > 0$, plugging $e^*_L(a)$ into $\frac{\partial u_H}{\partial e}(e,a)$ gives:

$$\text{(A6)} \qquad \frac{\partial u_H}{\partial e}(e^*_L, a) = w_H \frac{\partial h}{\partial e}(e^*_L, a) - \frac{\partial c}{\partial e}(e^*_L, a) > 0$$

Since $\frac{\partial u_H}{\partial e}(e, a)$ decreases with e, $e_H^*(a)$ should be larger than $e_L^*(a)$ to make Equation equal to 0.

Then consider changes of $e_\theta^*(a)$ with a. Note that $e_\theta^*(a)$ satisfies the FOC:

$$\frac{\partial u_\theta}{\partial e}(e_\theta^*, a) = w_\theta \frac{\partial h}{\partial e}(e_\theta^*, a) - \frac{\partial c}{\partial e}(e_\theta^*, a) = 0 \tag{A7}$$

Also note that the cross derivative is:

$$\frac{\partial^2 u_\theta}{\partial a \partial e}(e, a) = w_\theta \frac{\partial^2 h}{\partial a \partial e}(e, a) - \frac{\partial^2 c}{\partial a \partial e}(e, a) > 0 \tag{A8}$$

Equation (A8) is positive since $\frac{\partial^2 h}{\partial a \partial e}(e, a) \geq 0$ and $\frac{\partial^2 c}{\partial a \partial e}(e, a) < 0$. It suggests $\frac{\partial u_\theta}{\partial e}(e, a)$ increases with a. The worker's marginal net payoff of education e increases with ability a. As ability a increases to $a' > a$, $\frac{\partial u_\theta}{\partial e}(e_\theta^*(a), a') > \frac{\partial u_\theta}{\partial e}(e_\theta^*(a), a) = 0$. Again since $\frac{\partial u_\theta}{\partial e}(e, a)$ decreases with e, $e_\theta^*(a') > e_\theta^*(a)$ so that $\frac{\partial u_\theta}{\partial e}(e_\theta^*(a'), a') = 0$. Therefore, $e_\theta^*(a)$ increases with a, which means the optimal education for either L jobs or H jobs increases with ability a.

A3. Proof of Proposition 2

We first show $\underset{e}{max}\ u_H(e, a) \geq \underset{e}{max}\ u_L(e, a)$, so that all workers have incentives to receive their $e_H^*(a)$ education and get an H job if there is no constraint. It can be derived by the definition of maximum value:

$$\begin{aligned}
\underset{e}{max}\ u_H(e, a) &= u_H(e_H^*(a), a) \\
&\geq u_H(e_L^*(a), a)\ (e_H^*(a) \text{ gives the maximum}) \\
&= w_H h(e_L^*(a), a) - c(e_L^*(a), a) \\
&\geq w_L h(e_L^*(a), a) - c(e_L^*(a), a) \\
&= u_L(e_L^*(a), a) \\
&= \underset{e}{max}\ u_L(e, a)
\end{aligned} \tag{A9}$$

Next, we show the difference between $\underset{e}{max}\ u_H(e, a)$ and $\underset{e}{max}\ u_L(e, a)$ increases with the ability level a. Define the difference $\Delta(a)$ as:

$$\begin{aligned}
\Delta(a) &= \underset{e}{max}\ u_H(e, a) - \underset{e}{max}\ u_L(e, a) \\
&= u_H(e_H^*(a), a) - u_L(e_L^*(a), a) \\
&= [w_H h(e_H^*(a), a) - c(e_H^*(a), a)] - [w_L h(e_L^*(a), a) - c(e_L^*(a), a)]
\end{aligned}$$

Then consider the first-order derivative of $\Delta(a)$:

$$\begin{aligned}\Delta'(a) = {} & w_H \frac{\partial h}{\partial e}(e_H^*(a), a) \cdot \frac{\mathrm{d}e_H^*(a)}{\mathrm{d}a} + w_H \frac{\partial h}{\partial a}(e_H^*, a) \\ & - \frac{\partial c}{\partial e}(e_H^*(a), a) \cdot \frac{\mathrm{d}e_H^*(a)}{\mathrm{d}a} - \frac{\partial c}{\partial a}(e_H^*(a), a) \\ & - w_L \frac{\partial h}{\partial e}(e_L^*(a), a) \cdot \frac{\mathrm{d}e_L^*(a)}{\mathrm{d}a} - w_L \frac{\partial h}{\partial a}(e_L^*, a) \\ & + \frac{\partial c}{\partial e}(e_L^*(a), a) \cdot \frac{\mathrm{d}e_L^*(a)}{\mathrm{d}a} + \frac{\partial c}{\partial a}(e_L^*(a), a)\end{aligned} \tag{A10}$$

Note that in the derivation of $e_H^*(a)$ and $e_L^*(a)$, the FOCs are:

$$w_\theta \frac{\partial h}{\partial e}(e_\theta^*, a) = \frac{\partial c}{\partial e}(e_\theta^*, a) \tag{A11}$$

Plugging Equation (A11) into (A10) gives:

$$\Delta'(a) = \underbrace{[\frac{\partial c}{\partial e}(e_H^*(a), a) - \frac{\partial c}{\partial e}(e_L^*(a), a)]}_{>0} - \underbrace{[\frac{\partial c}{\partial a}(e_H^*(a), a) - \frac{\partial c}{\partial a}(e_L^*(a), a)]}_{<0} \tag{A12}$$

Since $\frac{\partial^2 c}{\partial e^2}(e, a) > 0$, $\frac{\partial c}{\partial e}(e, a)$ increases with education e. Similarly, $\frac{\partial c}{\partial a}(e, a)$ decreases with education e since $\frac{\partial^c}{\partial e \partial a}(e, a) < 0$. Note that $e_H^*(a) > e_L^*(a)$, so $\frac{\partial c}{\partial e}(e_H^*(a), a) - \frac{\partial c}{\partial e}(e_L^*(a), a) > 0$, while $\frac{\partial c}{\partial a}(e_H^*(a), a) - \frac{\partial c}{\partial a}(e_L^*(a), a) < 0$.

Therefore, $\Delta'(a) > 0$, so the difference $\Delta(a) = \underset{e}{max}\ u_H(e, a) - \underset{e}{max}\ u_L(e, a)$ increases with ability a.

A4. Proof of Proposition 3

The breakeven condition for a worker with ability a is given by Equation (A13) below. Note that the RHS is fixed for given ability a. WOLG we consider the upper breakeven education $\overline{e}_b(a)$, since the discussion for the lower breakeven education $\underline{e}_b(a)$ is simialr.

$$u_H(e, a) = w_H h(e, a) - c(e, a) = \underset{e}{max}\ u_L(e, a) = u_L(e_L^*(a), a) - c(e_L^*(a), a) \tag{A13}$$

First note that $u_H(e_H^*(a), a) = w_H h(e_H^*, a) - c(e_H^*, a) > RHS$. Since $u_H(e, a)$ decreases with education e for $e > e_H^*(a)$ and $\lim_{e \to \infty} u_H(e, a) < 0$, by the Intermediate Value Theorem, there exists $\overline{e}_b(a) > e_H^*$ that satisfies Equation (A13).

We also want to show $\overline{e}_b(a)$ increases with ability a, i.e., the worker can afford more education e as her ability a increases. Rerrange Equation (A13) and define:

$$A(e,a) = [w_H h(e,a) - c(e,a)] - [u_L(e_L^*(a),a) - c(e_L^*(a),a)] \tag{A14}$$

Then consider how $A(e,a)$ changes with education e and ability a by taking the partial derivatives.

$$\frac{\partial A}{\partial e}(e,a) = w_H h'(e) - \frac{\partial c}{\partial e}(e,a) = \frac{\partial u_H}{\partial e}(e,a) < 0 \tag{A15}$$

Equation (A15) is negative since we are considering $\overline{e}_b(a) > e_H^*(a)$. Therefore, $A(e,a)$ decreases with education e.

(A16)

$$\begin{aligned}\frac{\partial A}{\partial a}(e,a) &= -\frac{\partial c}{\partial a}(e,a) - w_L h'(e_L^*(a)) \cdot \frac{\mathrm{d}e_L^*(a)}{\mathrm{d}a} + \frac{\partial c}{\partial e}(e_L^*(a),a) \cdot \frac{\mathrm{d}e_L^*(a)}{\mathrm{d}a} + \frac{\partial c}{\partial a}(e_L^*(a),a) \\ &= -\frac{\partial c}{\partial a}(e,a) + \frac{\partial c}{\partial a}(e_L^*(a),a)\end{aligned}$$

The second line again follows from the FOC of the derivation of $e_L^*(a)$. Note that $\frac{\partial^2 c}{\partial e \partial a} < 0$, so $\frac{\partial c}{\partial a}(e,a)$ decreases with educaton e. Since we are considering $e > e_H^*(a) > e_L^*(a)$, $\frac{\partial c}{\partial a}(e_L^*(a),a) - \frac{\partial c}{\partial a}(e,a) > 0$. Hence, $A(e,a)$ increases with ability a.

To investigate how $\overline{e}_b(a)$ changes with ability a, consider an increase of a to $a' > a$. Then $A(\overline{e}_b(a),a') > A(\overline{e}_b(a),a) = 0$. To make $A(\overline{e}_b(a'),a') = 0$ again, we need to have $\overline{e}_b(a') > \overline{e}_b(a)$. Therefore, $\overline{e}_b(a)$ increases with ability a.

The analysis for $\underline{e}_b(a)$ is similar, although we should also consider the extra initial condition. If the worker's net payoff from H jobs with zero education is still larger than the best L job:

$$A(0,a) = [w_H h(0,a) - c(0,a)] - [u_L(e_L^*(a),a) - c(e_L^*(a),a)] \geq 0 \tag{A17}$$

Then $\underline{e}_b(a) = 0$ since $A(e,a)$ increases with education e when $e < e_H^*(a)$. In this case, there is only one real breakeven education, which is $\overline{e}_b(a)$. The worker will desire an H job for any education below it.

When Equation (A17) is negative, however, the analysis for $\underline{e}_b(a)$ is quite similar to that of $\overline{e}_b(a)$. Using the same procedures, we can show that $\underline{e}_b(a)$ also exists and decreases with ability a. A worker with higher ability a can earn the maximum L job net payoff from an H job more easily.

A5. Proof of Proposition 4

Consider firm H's value from employing a worker with ability and education (a, e). Firm H is willing to employ a worker if the worker generates positive net payoff to the firm:

$$V(e, a) = g(a, h(e)) - w_H h(e) \geq 0 \tag{A18}$$

First we know $V(0, a) < 0$ since $g(a, h(0)) = 0$. To investigate the existence and properties of the education requirement, consider the first-order derivative of Equation (A18) with respect to education e:

$$\frac{\partial V}{\partial e}(e, a) = [\frac{\partial g}{\partial h}(a, h(e)) - w_H]h'(e) \tag{A19}$$

Since $h'(e) > 0$, the sign of $\frac{\partial V}{\partial e}(e, a)$ depends on $\frac{\partial g}{\partial h}(a, h(e)) - w_H$. Note that $\frac{\partial^2 g}{\partial h^2}(a, h) < 0$, so $\frac{\partial g}{\partial h}(a, h(e)) - w_H$ decreases with education e. Then there can be 3 cases:

1) $\frac{\partial g}{\partial h}(a, h(0)) - w_H \leq 0$.

   The marginal product of skills at the lowest level (which is also the highest marginal product) is lower than the labor cost w_H. Then $\frac{\partial V}{\partial e}(e, a) \leq 0$ for all education levels e, so $V(e, a)$ decreases with education e. Since $V(0, a) < 0$, firm H will not hire this worker at any education level. The education requirement range is empty for such ability a.

2) $\frac{\partial g}{\partial h}(a, h(e)) - w_H > 0$ for all education e.

   Although $\frac{\partial g}{\partial h}(a, h(e))$ decreases with education e, w_H is a lower bound of it. In this case, $V(e, a)$ increases with education e, so there exists an education level $\underline{e}_r(a)$ such that $V(\underline{e}_r(a), a) = 0$. This education $\underline{e}_r(a)$ is the minimum education requirement.

3) $\frac{\partial g}{\partial h}(a, h(e)) - w_H > 0$ for small e and $\frac{\partial g}{\partial h}(a, h(e)) - w_H < 0$ for large e.

   In this case, $V(e, a)$ has a maximum value at $\frac{\partial g}{\partial h}(a, h(\hat{e})) - w_H = 0$. If the maximum value $V(\hat{e}, a) < 0$, then firm H will not hire this worker, since this worker cannot generate positive profit. However, the maximum value $V(\hat{e}, a) > 0$, then firm H will hire the worker within a range of education e, and this gives the range of education requirement $(\underline{e}_r(a), \overline{e}_r(a))$.

Case (2) and case (3) are more interesting. We show that the minimum education requirement $\underline{e}_r(a)$ decreases with ability a. Similar analysis can be applied to the maximum education requirement $\overline{e}_r(a)$ and show it increases with ability a.

Consider the boundary condition:

$$V(\underline{e}_r(a), a) = g(a, h(\underline{e}_r(a))) - w_H h(\underline{e}_r(a)) = 0 \quad \text{(A20)}$$

The first-order derivative of $V(e, a)$ with respect to a is:

$$\frac{\partial V}{\partial a}(e, a) = \frac{\partial g}{\partial a}(a, h(e)) > 0 \quad \text{(A21)}$$

Equation (A21) suggests that $V(e, a)$ increases with ability. Thus, when ability increases from a to a', the boundary conditio becomes:

$$V(\underline{e}_r(a), a') = g(a', h(\underline{e}_r(a))) - w_H h(\underline{e}_r(a)) > 0 \quad \text{(A22)}$$

Since we are considering $\underline{e}_r(a)$ and $V(e, a)$ increases with e in its neighborhood, to make the boundary condition tight again, we need to have $\underline{e}_r(a') < \underline{e}_r(a)$.