Federal Reserve Bank of Cleveland Working Paper Series

Quantifying the Signaling Role of Education

Barış Kaymak

Working Paper No. 25-02

January 2025

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment on research in progress. They may not have been subject to the formal editorial review accorded official Federal Reserve Bank of Cleveland publications.

See more working papers at: www.clevelandfed.org/research. Subscribe to email alerts to be notified when a new working paper is posted at: https://www.clevelandfed.org/subscriptions.

Quantifying the Signaling Role of Education[*]

Barış Kaymak[†]

Federal Reserve Bank of Cleveland

Abstract

This paper quantifies the signaling role of education and measures the associated efficiency losses from asymmetric information. To that end, I model educational attainment and occupational choices in an asymmetric information environment with employer learning and socially productive education. The model highlights how occupational sorting and the pace of employer learning jointly determine the strength of the signaling motive in equilibrium. I estimate the signaling role of education versus human capital by relating differences in employer learning across occupations, which generates variation in signaling incentives, to the distribution of ability and educational attainment by occupation. The estimates suggest that the role of job market signaling relative to the human capital model is 23 percent. On the margin, a year of additional schooling raises productivity by 6.4 percent and the return to signaling is 2.4 percent. In a counterfactual analysis, eliminating asymmetric information reallocates labor from education to workforce participation and improves occupational sorting. Aggregate efficiency gains are equivalent to 7.6 percent of lifetime earnings.

1   Introduction

In a model of job market signaling, individuals invest in education not only to acquire valuable job skills but also to signal their ability to otherwise uninformed employers (Spence, 1973). Hence, the wage return to education captures both the value of acquired skills and the value of perceived ability by potential employers. Since the signaling component does not raise productivity but merely its perception, educational attainment in such models is inefficiently high. This paper estimates the return to the signaling role of education and measures the associated efficiency losses.

The key challenge in identifying the signaling model lies in distinguishing it from a human capital model, since both predict similar equilibrium behavior. In both models, equilibrium wages reflect expected productivity, and workers choose their education given the relationship between education and wages. This makes it difficult to ascertain the relevance of the signaling motive for schooling decisions.

I address this challenge by modeling educational and occupational choices in an asymmetric information setting with employer learning. The empirical strategy exploits variation in employer learning across occupations, which creates differences in signaling incentives. In occupations where productivity is easily revealed—either through interviews or on-the-job performance—education plays a limited signaling role, resulting in lower educational attainment for a given ability level.

An important hurdle for this strategy is that occupations might differ not only in signaling returns but also in how education augments productivity. If education improves productivity more in occupations where worker performance is harder to evaluate, one could falsely attribute a larger role to signaling. I overcome this hurdle by allowing education's contribution to productivity to differ across occupations and by exploiting the opposing sorting patterns predicted by human capital and signaling theories. I show that while workers with higher ability have a comparative advantage in sectors where human capital is more valuable, workers with lower ability have a comparative advantage where signaling is more important. This occurs because the signaling equilibrium imposes a heavier burden on high-ability workers, who must invest substantially in education to distinguish themselves from low-ability workers. Low-ability workers, who in equilibrium achieve the lowest education levels regardless of the information environment, face lower costs from signaling. These contrasting predictions allow me to distinguish the relative importance of human capital versus signaling by examining ability sorting patterns in combination with the patterns of employer learning and wage returns to education across occupations.

Using data on the distribution of ability and education across occupational groups, I find that observed educational attainment is consistent with a hybrid model where signaling accounts

for 23 percent and human capital for 77 percent of returns to education. An additional year of education raises productivity by 6.4 percent and provides a signaling return of 2.4 percent, yielding a total private return of 8.8 percent.

Across the ability distribution, uncertainty about worker productivity resolves early (partly at the recruitment stage) in occupations that attract high-ability workers, limiting signaling incentives. By contrast, low-ability occupations display a delayed resolution. Because signaling costs rise with ability, this sorting pattern, which is consistent with the signaling theory, suggests that self-selection limits the efficiency losses from asymmetric information. In a counterfactual analysis where asymmetric information is eliminated, the resulting reallocation from education to workforce participation generates efficiency gains of 7.6 percent of lifetime output for an average worker.

This paper advances the literature on job market signaling in a couple of dimensions. First, it quantifies the wage return to job market signaling and provides an estimate for the efficiency losses associated with asymmetric information in terms of forgone output. Second, it extends the standard job market signaling model (Spence, 1973) to multiple markets and integrates it with the literature on statistical discrimination under employer learning (Farber and Gibbons, 1996; Altonji and Pierret, 2001). These extensions result in a hybrid model of education, where the overall strength of the signaling motive depends not only on how quickly the uncertainty resolves but also on the occupational choices of workers. Both are important for aggregate efficiency losses from signaling.

Earlier studies tested for job market signaling by comparing educational choices in markets where signaling incentives presumably differ in strength, such as the self-employed versus salaried workers (Wolpin, 1977), or occupational and industrial categories (Riley, 1979a). Similarly, Lang and Kropp (1986) and Bedard (2001) compare the distribution of educational attainment before and after a change in the institutional environment to test for signaling. While the findings in this literature attribute a partial role to signaling, they do not measure its magnitude nor its efficiency implications.

Subsequent work recognized the importance of employer learning for the signaling role of education (Altonji and Pierret, 1997; Lange, 2007). Lange (2007), in particular, estimates the speed of employer learning and provides an upper bound on the return to signaling using the optimality condition for schooling choice under asymmetric information in a single-sector model. Nonetheless, whether educational attainment respects this optimality condition, i.e., whether agents do in fact behave strategically in their schooling choices, remains unclear.[1] Fang (2006) estimates a

---

[1]The employer learning literature typically does not distinguish between symmetric uncertainty, where workers learn their ability alongside their employers, and asymmetric uncertainty, where workers know their ability but employers do not. The signaling incentive is only present in the latter case.

static model of college choice with ability signaling using data on the distribution of wages and college attainment for young workers and finds that signaling explains about a third of the college wage premium at labor market entry.[2] While similar in its imposition of modeling restrictions on the data, this paper differs by examining broad occupation groups that provide different signaling incentives, and by establishing the consistency of educational choices and occupational sorting patterns with the theory of signaling.[3]

Conceptually, this paper is also related to Riley (1979a), who compares differences in educational attainment by occupation to test for signaling. The analysis here uses estimates of employer learning to distinguish occupations that are potentially subject to signaling, compares educational attainment *conditional on ability* as proxied by cognitive test scores, and allows for non-trivial sorting into occupations by ability.[4]

The rest of the paper is organized as follows. The next section presents the model and outlines the roles of employer learning and occupational choices for the signaling role of education. Section 3 describes the empirical strategy and Section 4 presents the estimation results. Section 5 concludes.

2   A Model of Signaling with Employer Learning and Market Choice

The objective of this section is to model schooling choice under asymmetric information when there is employer learning. I start with a single-sector model and analyze the interaction between employer learning and signaling. Section 2.1 extends the model to multiple sectors and provides an analysis of occupational choice under asymmetric information.

The economy is populated by a continuum of infinitely lived workers with heterogeneous skills. The marginal product of a worker with education $s_i \geq 0$ and ability $a_i \geq \underline{a}$ is given by[5]

$$\ln y_{it} = q(s_i, a_i) + h_t + \varepsilon_i + \nu_{it},$$

where $q(.)$ is increasing in education and ability and strictly concave in education ($q_s, q_a \geq$

---

[2]Fang (2006) also reports an estimate for workers with 21-24 years of experience, where signaling is 25 percent of the college wage premium. This is at odds with the estimates of employer learning here and in Lange (2007), which suggest that most uncertainty is resolved early, driving the return to signaling to nearly zero for experienced workers.

[3]Altonji (2005) and Gibbons et al. (2005) study occupational sorting in a human capital model with employer learning under symmetric uncertainty. I consider an asymmetric information setting instead, where occupations differ in their valuation of not only human capital but also ability signaling.

[4]In Riley (1979a), compensating differentials ensure that agents are indifferent between different occupations, so there is no reason to expect differences in ability across occupations.

[5]I assume that ability is bounded from below. This ensures the uniqueness of the informational equilibrium.

4

$0, q_{ss} < 0$).[6] Hence, education raises productivity regardless of its signaling value. The marginal return to education is strictly increasing in ability, $q_{sa} > 0$. Following the employer learning literature (Altonji and Pierret, 2001; Farber and Gibbons, 1996; Lange, 2007), I assume that wages increase with experience in a deterministic way described by h_t. The worker-specific random component ε_i is independent of a_i and s_i.[7] One could think of a_i as cognitive ability and ε_i as market ability. Workers are informed of their a_i, but are uninformed about ε_i. Neither component is directly observable to employers. Finally, ν_{it} is a stochastic productivity term that is not observed by either party. For tractability, I assume that ε_i and ν_{it} are normally distributed with 0 means, and variances $\sigma_{\varepsilon_i}^2$ and σ_ν^2.

At the onset of a worker's career, two pieces of information are available to employers: their education and their application package. The application package is an umbrella term for additional information extracted by the employer from an applicant's résumé, references, job interview etc., during the recruitment process. These alternative screening tools limit the signaling role of education and constitute the first component of the employer learning process. To capture this component in a simple way, I assume that upon graduation, a worker's true productivity is observed fully with probability $\rho < 1$. Otherwise, his productivity is revealed slowly over time. In what follows I set $\rho = 0$ for simplicity as the value of ρ has no bearing on the qualitative features of the model. Later, I relax this assumption to estimate the model.

Once a worker starts his career, employers observe a noisy measure of his output every period.

$$\eta_{it} = \ln y_i + u_{it} \tag{1}$$

The noise term u_{it} is normally distributed with mean 0 and variance σ_u^2. Let $I_{it} = \{\eta_{i1}, \eta_{i2}, ..., \eta_{it}\}$ denote the history of worker i's observed output up to experience level t. By construction, $I_{i0} = \emptyset$. Signals of output are observed by all employers in a competitive market.

Workers maximize their lifetime income by choosing their education level conditional on their information set. For simplicity, workers are assumed to have infinitely long careers. This assumption is relaxed below when the model is estimated.

$$V(a_i; \Lambda) = \max_s \sum_s^\infty \delta^{t+s} E[W_{it}(s)|s_i, a_i], \tag{2}$$

where $W_{it}(s)$ denotes the wage offer to a worker with s years of education and t years of experience, and δ is the discount factor. The cost of education is determined by forgone earnings during the time spent at school.

A signaling equilibrium in this simple setting is defined as follows.

---

[6]Strictly diminishing returns are not necessary when career length is finite.
[7]This is without loss of generality since any correlation can be embedded in $q(.)$

Definition 1. *(Riley, 1979a) A signaling equilibrium is a wage function $W_{it}(s)$ and a policy function $S(a)$ such that,*

1. *For all $a_i \geq \underline{a}$, $S(a_i)$ solves (2) given $W_{it}(s)$*

2. *Individual choices are consistent with the wage function,*

$$W_{it}(s) = E[y_i | s_i, I_{it}]$$

The second condition follows from the assumption that the labor market is competitive. The existence of an equilibrium wage function with informational consistency is discussed in detail in Riley (1979a). There are essentially two requirements. First, in the economy without asymmetric information, where productivity is directly observed, there is a unique education level that solves (2) for each ability level a. Second, the opportunity cost of attaining higher educational levels is lower for workers with higher ability. The former is ensured by the concavity of $q(s, a)$ in s, and the latter is met by the complementarity between cognitive ability and education in human capital production: $q_{sa} > 0$. A separating equilibrium where workers with higher a choose higher education thus exists.

I now turn to the characterization of the equilibrium. Denote the ability level inferred by the market given an education level s by the function $A(s) : \mathbb{R}_+ \to [\underline{a}, \infty)$. Upon observing an output signal η_{it}, the employer extracts noisy information on ε_i by $\eta_{it} - q(s_i, A(s_i)) = q(s, a_i) - q(s, A(s_i)) + \varepsilon_i + u_i + \nu_{it}$. Note that these are unbiased observations around the true ε_i if and only if the inferred ability $A(s_i)$ equals the true ability a_i, as in a perfectly separating equilibrium. The wage offered to a worker is the expected productivity conditional on his education level s_i and the history of signals I_{it}. Let $\tilde{\sigma}_y = \sigma_\varepsilon^2 + \sigma_\nu^2$ be the variance of wages conditional on education and ability. Under the distributional assumptions for ε_i, u_i and ν_{it}, wages can be expressed as:

$$\ln W_{it}(s_i, I_{it}) = (1 - \lambda_t) q(s, A(s)) + \lambda_t \bar{\eta}_{it} + \tilde{h}_t, \tag{3}$$

where

$$\lambda_t = \frac{\tilde{\sigma}_y^2}{\tilde{\sigma}_y^2 + \sigma_u^2/t}, \quad \bar{\eta}_{it} = \sum_{k=1}^{t} \eta_{ik}/t, \quad \text{and} \quad \tilde{h}_t = h_t + 0.5(1 - \lambda_t^2)\tilde{\sigma}_y^2.$$

A worker's wage is a weighted average of his productivity, as perceived by the market conditional on education alone, and the average observed output up to t. Employers rely less on education and more on average output as λ_t increases from $\lambda_0 = 0$ to $\lambda_\infty = 1$.

Workers base their education decisions on the discounted value of expected wages in the future conditional on their own ability. Given equation (3), the expected wage stream at the beginning of a worker's career is

$$\ln E_0[W_{it}(s_i, I_{it})|s_i, a_i] = q(s, A(s))(1 - \lambda_t) + q(s, a)\lambda_t + h_t + \tilde{\sigma}_y^2/2. \tag{4}$$

Plugging equation (4) into (2), the maximization problem is equivalent to[8]

$$\ln V(a_i; \Lambda) = \max_s \ \Lambda q(s, A(s)) + (1 - \Lambda)q(s, a_i) - rs, \tag{5}$$

where

$$\Lambda = \frac{\sum_0^\infty (1 - \lambda_t)e^{-rt+h_t+\tilde{\sigma}_y^2/2}}{\sum_0^\infty e^{-rt+h_t+\tilde{\sigma}_y^2/2}}, \tag{6}$$

and $r = -\ln\delta$ is the discount rate. The objective function described above is a convex combination of a typical signaling model, and a typical human capital model. The term $\Lambda \in [0, 1]$ measures the extent of signaling in the model. When the output signals are completely uninformative, $\sigma_u \to \infty$, then employers rely only on education to estimate productivity at all times, $\lambda_t = 0$ for all t, and the model reduces to a typical signaling model: $\Lambda = 1$. On the other extreme, if output is fully observed, then employers do not rely on education at all, $\lambda_t = 1$ for all t, and the model becomes a pure human capital model: $\Lambda = 0$.

In the more general version of the model, where the uncertainty about productivity partially unravels during the recruitment process, $\rho > 0$, Λ is replaced by $\tilde{\Lambda} = (1 - \rho)\Lambda$. The qualitative features of the model are thus maintained as long as $\rho < 1$, i.e., some uncertainty remains.

The optimality condition for the schooling choice is

$$\Lambda \left(q_s(s, A(s)) + q_a(s, A(s))A'(s) \right) + (1 - \Lambda)q_s(s, a_i) = r,$$

which defines a one-to-one mapping from the ability level a_i to the schooling choice s. Therefore, at a separating equilibrium the schooling choice fully reveals a worker's ability, implying $A(S(a_i)) = a_i$ for all a_i. Using this equality, the first order condition simplifies to

$$q_s(s, a_i) + q_a(s, a_i)A'(s)\Lambda = r. \tag{7}$$

At the equilibrium, workers equate the private return to education to the discount rate. The first term on the left side of equation (7) represents the increase in productivity due to higher human capital. The second term represents the marginal return to signaling: an increase in education

---

[8]Constant terms that do not depend on s or a are dropped from the objective function.

raises employers' perception of ability by $A'(s)$ units on the margin, which is priced at $q_a(s_i, a_i)$ per unit in the market. Since true productivity is eventually revealed by job performance, this component is limited by $\Lambda \in [0, 1]$.

Note that even though employers are trying to learn ε, a variable that is not related to education or ability, schooling choices are affected by the pace of learning. This is because the information on ε that is inferred from observed output relies on the accuracy of employers' prior about ability. The latter gives the worker the incentive to try to delude the employer by attaining a higher level of education, but, of course, the employers' predictions are correct at the (separating) equilibrium.

Equation (7) defines a differential equation in s, the solution to which describes the equilibrium strategy for employers. Without an initial condition, the solution to this problem is a continuum of functions that differ in their intercepts. The following lemma establishes that the worker with minimum ability chooses his schooling level efficiently, pinning down a unique equilibrium strategy.

Lemma 1. *Let $S^*(a) = \{s \geq 0 : q_s(s, a) = r\}$ be the socially efficient level of education. For any $\Lambda \in [0, 1]$, $S(\underline{a}; \Lambda) = S^*(\underline{a})$ at a separating equilibrium.*

Proof. $S(0) < S^*(0)$ clearly cannot be the case, since the agent can still signal his ability and improve his lifetime income by receiving $S^*(\underline{a})$. Suppose for a contradiction that $S(\underline{a}) > S^*(\underline{a})$. At a separating equilibrium we must have $A(S(\underline{a})) = \underline{a}$. Consider a deviation to $S^*(\underline{a})$, for which $A(S^*(\underline{a})) = \underline{a}$ at the equilibrium, since it is still the minimum of all education choices. But then, $V(S(\underline{a}), \underline{a}) = e^{-rS(\underline{a})} \int_{\underline{a}}^{\infty} e^{-rt} \exp q(S(\underline{a}), \underline{a}) dt < e^{-rS^*(\underline{a})} \int_{\underline{a}}^{\infty} e^{-rt} \exp q(S^*(\underline{a}), \underline{a}) dt = V(S^*(\underline{a}), \underline{a})$ by definition of $S^*(\underline{a})$, which contradicts the optimality of $S(\underline{a})$ in the signaling environment. $\square$

With this lemma, the signaling equilibrium with employer learning can be fully characterized.

Proposition 1. *The fully separating equilibrium is characterized by functions, $S(a)$ and $A(s)$, such that*

1. *$A(s)$ solves the differential equation in (7) with the initial condition $A(S^*(\underline{a})) = \underline{a}$.*

2. *Schooling choices are consistent with the market's predictions.*

$$S^{-1}(a) = A(s)$$

3. *Given random components ε_i and $\{u_{it}\}_{t=1}^{\infty}$, worker i's wage process is given by,*

$$\ln W_{it} = q(s_i, a_i) + \lambda_t(\varepsilon_i + \bar{u}_{it}) + \tilde{h}(t)$$

where $\bar{u}_t = \sum_{k=1}^{t} u_{ik}/t$

The last part of the proposition can be obtained by setting $A(s_i) = a_i$ and using equation (1) in the wage equation.

Part 3 of the proposition highlights the infamous identification problem in the empirical literature on signaling. The returns to education and ability observed in the wage data do not depend on the extent of signaling, Λ. At a separating equilibrium, employers observe education and correctly infer ability. The wages, therefore, correctly reflect the marginal returns to education and ability, given the schooling choices.

In most applications, the econometrician does not observe ability. In this case, the estimated return to education captures not only the increase in productivity (often referred to as the social return), but also the differences in the abilities of workers with different education levels, known as the ability bias. The private return to education lies between the estimated return and the social return. To see this, consider the total derivative of the wage function with respect to education:

$$\frac{d\mathbb{E}[\ln W_{it}|s]}{ds} = \overbrace{\underbrace{\frac{\partial \ln q(s,a)}{\partial s}}_{Social\,Return} + \Lambda \frac{\partial \ln q(s,a)}{\partial a} A'(s)}^{Private\,Return} + \underbrace{(1-\Lambda)\frac{\partial \ln q(s,a)}{\partial a} A'(s)}_{Ability\,Bias}, \tag{8}$$

where the ability bias is conveniently decomposed into two by $\Lambda \in [0,1]$. The first term is the marginal increase in productivity and the second term is the return to signaling. Together, they constitute the private return to education, which is equated to the discount rate at the optimum (see equation (7)). The sum of the last two terms represent the ability bias, which is independent of the value of Λ, given the schooling choices. Therefore, measuring signaling requires not the elimination of the ability bias, but, instead, a decomposition of it.

While $\Lambda \in [0,1]$ measures the relative importance of the signaling model, the quantitative significance of the return to signaling, the second term in (7), also depends on the significance of ability for productivity, $q_a()$, and the slope of the ability-education gradient, $A'(s)$. These components are generally functions of the economic environment as well as the severity of asymmetric information. Consequently, the magnitude of the return to signaling relative to the social return to education can be quite different than Λ.

The impact of the degree of signaling on the wage distribution operates through its effect on educational choices. A larger role for signaling increases the marginal return to schooling investment and raises educational outcomes for a given ability level:

Proposition 2. *For each ability type a, the schooling level at the separating equilibrium $S(a; \Lambda)$ is increasing in the degree of signaling, Λ.*

In combination with Lemma 1, Proposition 2 suggests that agents with higher abilities are affected more severely by signaling. Since the educational choice of the lowest-ability individuals is always fixed at the efficient level, while the educational choices of higher-ability individuals rise with Λ, the education ability profile becomes steeper with Λ.

A steeper education-ability profile in environments with a higher degree of signaling reduces the inferred ability level by a given level of education, and implies a lower return to signaling, $A'(s)$. This implies that markets with a larger role for signaling are characterized by a *smaller* ability bias, a point made by Lang (1994). Moreover, the additional investment in education in a market with more signaling reduces the return to education on the margin due to diminishing returns, which leads to a lower overall estimate.

The model also allows for a simple expression for the bound on signaling proposed by Lange (2007). To see this, denote the total return to education that contains the true productivity-augmenting effect and the ability bias by $b_s = q_s(s, a) + q_a(s, a)A'(s)$. Substituting b_s in the first-order condition for education choice given by (7), and rearranging terms gives the following expression for the return to signaling.

$$\Lambda q_a(s, a)A'(s) = \frac{\Lambda}{1 - \Lambda}(b_s - r) \tag{9}$$

Under linearity assumptions, b_s can be estimated by a simple regression of wages on education without controlling for ability. Given also an estimate of Λ, the return to signaling can be obtained up to the private return to education r. However, a consistent estimate of Λ requires full access to the employers' information set regarding each worker. Therefore, Lange (2007) considers his estimate of Λ as a lower bound for the true value. As a result, the estimator for the return to signaling suggested by equation (9) overestimates the true return to signaling.

The next section extends the model to multiple sectors with differing roles for signaling and human capital, and analyzes the market choice of income-maximizing agents.

2.1 Multiple Sectors and Market Choice

To study the endogenous market choice, I make two critical assumptions for tractability. First, workers must commit to a sector before entering the labor market. Once they are employed, the information flow is common knowledge for all employers in all sectors. This rules out systematic switching of sectors based on employer learning at the equilibrium. In the empirical section, I use data on expected future occupations along with actual occupation to gauge the implications

of this assumption for the results.

Second, I assume that workers are endowed with individual sector preferences that are uncorrelated with the fundamental variables in the model, such as education and ability. This prevents the market choice itself from being an informative signal for ability.[9] It also implies that all sectors have some workers of all ability levels at the equilibrium, but, on average, some sectors will have workers with higher ability.

Let $j \in \{1, 2\}$ denote sectors. The productivity of a worker is given by:

$$\ln y_{it} = \ln \omega_j + p_j q(s_i, a_i) + h_t + \varepsilon_i + \nu_{it},$$

where ω_j is the sector-specific price of an efficiency unit of human capital and p_j denotes the sector-specific return to human capital. The experience profile of productivity is assumed to be the same in both sectors.[10] Note that ε_i is valued equally in both sectors. This ensures that workers do not have an incentive to switch sectors based on the realizations of output signals because both the current wage and the expected wage in the alternative sector change similarly as information is common to the market.

A worker's earnings are determined by his expected productivity valued at the sector-specific price, ω_j: $W_{ijt} = \omega_j E[y_{ij}|j, s_i, I_{it}]$. Workers are endowed with preferences for market j, denoted by $\tau_{ij} \in (-\infty, +\infty)$. It is assumed that $\Delta\tau = \tau_{i1} - \tau_{i2}$ is independent of a_i, and is distributed according to the cdf $G(\Delta\tau)$. The sectors differ primarily in two dimensions: the signaling role of education, measured by Λ, and the return to human capital, denoted by p. The value of choosing sector j for a worker with ability a_i is:

$$\ln V^j(a_i, \tau_{ij}) = \max_s \ln \omega_j + \Lambda_j\, p_j q(s, A(s)) + (1 - \Lambda_j)\, p_j q(s, a_i) - rs + \tau_{ij} \qquad (10)$$

To establish the single crossing property that characterizes selection in this economy, consider the slope of the value function with respect to ability. The envelope condition implies

$$\frac{dV^j(a_i, \tau_{ij})}{da_i} = (1 - \Lambda_j) p_j q_a(s, a_i), \qquad (11)$$

which is increasing in p and decreasing in Λ. Workers with higher ability are more likely to sort into markets where human capital is rewarded well and the signaling role of education is small. To see the intuition behind the latter, recall that signaling is costly: it pushes workers beyond the optimal level of education, but, at the equilibrium, employers always correctly infer ability. Workers would, therefore, rather eliminate signaling altogether. This cost is particularly large

---

[9]Although sectoral choice as a signaling device is an interesting premise for further research, it considerably complicates the informational equilibrium in the current setting with little value added to the analysis.

[10]I relax this assumption in the empirical section.

for workers with the highest abilities since they need to invest the most in order to distinguish themselves from everyone else. The worker with the lowest ability, for instance, never suffers since he invests optimally in both markets.

The sector-specific prices, $\{\omega_j\}_{j=1,2}$, are determined endogenously by the equilibrium demand and supply in each sector. This prevents degenerate sorting behavior where all workers prefer the sector where human capital is more valuable or where signaling is not important. It also gives a more accurate prediction of occupational and educational choices in counterfactual simulations that are used to measure efficiency losses associated with signaling. Let $p_j(\Delta\tau) = \{a_i \in [\underline{a}, \infty) :$ $\ln V_j(a_i, \tau_{ij}) > \ln V_k(a_i, \tau_{ik}), j = 1, 2, \ k \neq j\}$ be the abilities of workers with sectoral preference $\Delta\tau$ for whom sector j is optimal. Equation (11) implies that $p_j(\Delta\tau)$ are intervals determined by ability levels that are below or above a threshold ability level $\tilde{a}(\Delta\tau)$. The total supply of human capital in each sector is:

$$ h_j = \int_{\Delta\tau} \int_{a \in p_j(\Delta\tau)} \exp\left(p_j q(S_j(a), a) + h_t + \sigma_{\tilde{y}}^2/2\right). $$

Production in each sector is linear in the stock of human capital in that sector, h_j, and total output is obtained by a Cobb-Douglas aggregator: $Q = zh_1^\phi h_2^{1-\phi}$. Assuming markets are competitive, relative sectoral factor prices are given by:

$$ \ln(\omega_1/\omega_2) = \ln(\phi/(1-\phi)) + \ln(h_2/h_1). \tag{12} $$

Equation (12) implies that the factor prices are consistent with the sectoral allocation of labor implied by the sorting behavior. It also ensures a non-degenerate sorting of workers into sectors. If all workers were to choose sector 1, then the relative wage rate per efficiency unit of human capital in sector 2 approaches infinity. A positive measure of workers then would find it optimal to switch to sector 2.

Figure 1 depicts the equilibrium sorting behavior between two sectors that differ in the strength of signaling, Λ_H and $\Lambda_L < \Lambda_H$, and the associated education choices in each sector. The straight lines in Panel (a) show the mean value function in each sector, obtained when $\tau_{ij=0}$. For each ability level, a worker compares $\ln V^L(a, 0)$ with $\ln V^H(a, 0)] + \Delta\tau$. Given the distribution of $\Delta\tau$, workers with lower ability are more likely to choose sector H and workers with higher ability are more likely to choose sector L. Panel (b) displays the education choices in each sector. Those in sector H attain higher levels of education. The slope of the education-ability profile is also steeper for this sector. The dashed curve depicts the perfect information case, where workers' education choices are socially efficient.

(a) Sector Decision   (b) Education Decision
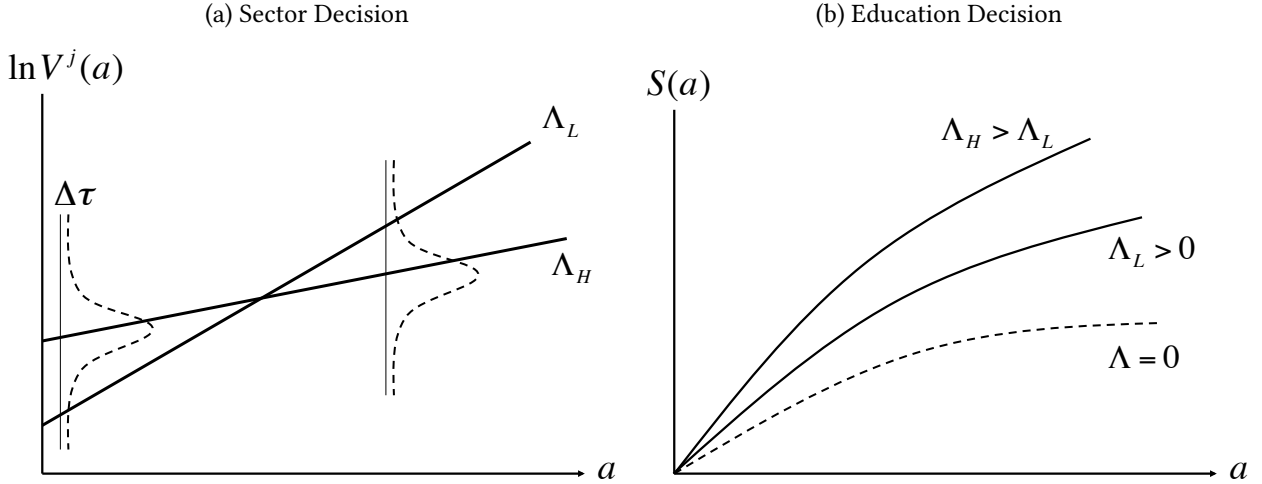
Note.– Panel (a) shows the value functions for two sectoral options that differ in the role of signaling, $\Lambda_L < \Lambda_H$. $\Delta\tau$ is the relative idiosyncratic preference for sector H. Panel (b) shows the associated education ability profile in each sector. Workers with higher ability are more likely to sort into sectors where the signaling role of education (Λ) is lower, and attain a lower level of education conditional on ability.

3   Quantifying the Model

The model predicts that markets with a larger role for signaling are characterized by higher educational attainment conditional on ability. Furthermore, schooling investment is more sensitive to differences in ability, implying a larger variance of education in such markets. When worker mobility is not restricted, workers with higher ability are concentrated in markets with a smaller role for signaling. These predictions allow for the estimation of the return to signaling by comparing the implications of the model for educational attainment and ability across markets with varying roles for signaling, defined here by occupations.

The choice of occupations as the relevant market relies on the presumption that the extent of asymmetric information differs by occupations. In particular, the pace of employer learning may be slower in some occupations due to production lags, or complex interactive tasks, e.g., teamwork, that are involved in the production process.[11] The strategy here is to estimate the potential roles for signaling in each category by estimating the speed of employer learning, and then testing if the joint distribution of education and ability is consistent with the signaling model given the potential roles for signaling.

The production function used in the estimations is $q(s, a) = p_j s^\beta a^\alpha$. This function satisfies the assumptions of the model and features equilibrium wage profiles that are log-linear in education,

---

[11]Riley (1979b) compares different occupations to test for signaling with a similar motivation.

a salient feature of the data. With this specification, the parameters of the model can be classified into four groups: the parameters that describe the distributions of ability and preferences (ε, a, τ), the environmental parameters (r, T), the production function parameters $(\alpha, \beta, \{p_j\}, \phi)$ and the signaling parameters, Λ_j, where j denotes occupational category. Some of the parameters are observed in the data, and the remaining ones are estimated indirectly using the simulated method of moments.

The estimation of the speed of employer learning is based on a version of the method introduced in Lange (2007) adapted here to suit non-linear models.[12] The procedure requires data on wages, educational attainment, and ability, as measured by test scores, for a panel of workers. Following Lange (2007) and Altonji and Pierret (2001), it is assumed that the test scores in the data are not directly observed by employers, but contain information on ability that is learned after a worker is hired.

Let z_i denote the test score, and define the residuals from a regression of the test score on educational attainment by $\tilde{z}_i = z_i - E[z_i|s_i]$. Since the model features perfect separation, we interpret the predictable component of the test score as the learning ability a in the model, and the residuals as ε. With this interpretation, the parameters that describe the distributions of a and ε are identified directly from the data. In particular, $\hat{\sigma}_\varepsilon^2 = var[\tilde{z}_i]$ and $\hat{\sigma}_{\ln a} = var(E[z|s])$, where we assume that a is distributed log-normally and ε is distributed normally. Similarly, the means are defined as $\mu_{\ln a} = E[z]$ and $\mu_\varepsilon = 0$.

To estimate the speed of learning, consider the following regression:

$$\ln w_{it} = q(s, A(s)) + \kappa_{zt}\tilde{z}_i + X_{it}\Gamma + \nu_{it}, \tag{13}$$

where log-wages are regressed on a function of educational attainment, the residual test score interacted with indicators of potential experience and other controls. If the test score residual, $\tilde{z}_i$, contains information that is not available to employers at the beginning, but is learned later on the job, then it becomes increasingly important over time. This suggests that the coefficient on $\tilde{z}$ increases with experience. The speed of learning can be measured by how fast κ_{zt} rises. Note that unlike in Lange (2007), the estimated relationship between wages and education, as approximated by the derivative of $q(s, A(s))$, does not change with experience. The reason is that the method here uses the residuals from the test score, which are, by construction, orthogonal to education.[13]

Given the equilibrium wage equation in (3), the coefficients on $\tilde{z}$ can be expressed as $\kappa_{zt} = \kappa_{z0}(1 - \lambda_t) + \lambda_t\kappa_{z\infty}$, where $\kappa_{z\infty}$ is the return to ability, and $\kappa_{z0} = (1 - \rho)\kappa_{z\infty}$ is the part of that

---

[12]The estimation in Lange (2007) is based on the learning model in Altonji and Pierret (2001), whereas the estimation here builds on a similar model by Farber and Gibbons (1996).

[13]Theoretically, this is because employers' predictions regarding ability are correct on average. Therefore, updates to inferred ability are zero on average.

return that is captured by alternative signals of ability available to employers during the recruitment process. The pace of learning that occurs on the job is captured by λ_t, which increases from 0 to 1 with experience (see equation (3)). The speed of learning implicit in λ_t is determined by the signal to noise ratio $\kappa = \sigma_u^2 / \sigma_{\tilde{y}}^2$. A higher κ implies that λ_t increases slowly with experience, corresponding to a slow learning process. Given the coefficients on $\tilde{z}$, the parameters $\{\kappa_{z0}, \kappa, \kappa_{z\infty}\}$ can be estimated by the following non-linear regression.

$$\hat{\kappa}_{zt} = \kappa_{z0} \frac{\kappa}{t + \kappa} + \frac{t}{t + \kappa} \kappa_{z\infty}. \tag{14}$$

Once the signal to noise ratio, κ, is estimated, the role of signaling for each category, Λ_j, can be computed up to ρ_j. The true return to signaling could not be estimated this way because employers may have access to alternative signals that is not available to the researcher. If this information is correlated with the observed test score, it would show up in κ_{z0}, but otherwise, it would not be captured. Therefore, ρ_j are estimated indirectly using the implications of the model.

The length of life T is set to 54, representing ages from 6, when agents start school, to 60 when they retire. The remaining parameters of the model are the discount rate, r, the signaling parameters ρ_j, the standard deviation of the occupational taste parameter, σ_τ, and the production function parameters $(\alpha, \beta, \{p_j\}_{j=1,2}, \phi)$. These parameters are estimated using the model's predictions for the education profiles in the two occupations and the differences in the test scores. The particular moments we match are the average education, and the gradient of the education-ability profile in each sector, the mean and the standard deviation of ability in each sector and the share of workers in the high-ability sector.

Although it is difficult to link the identification of each parameter to a particular moment, an intuitive argument can be provided. The distributions of ability and education within occupations are governed by both human capital and signaling roles of education in that occupation. In principle, both components could explain the observed distributions. However, these components have different implications for the sorting into occupations and, therefore, the differences in the distribution of ability and education across occupations. Consider, for instance, a sector, I, where workers have a higher level of education relative to another sector, II. This can occur either because the signaling plays a major role, or because human capital is particularly valuable in sector I. If the signaling theory is the primary explanation, then the sorting behavior suggests that sector I is expected to contain high ability workers. Therefore, the degree of sorting by ability indicates the relative role of each component.

4 Data and Estimation Results

Data are taken from the NLSY, which contains results from the Armed Forces Qualification Test (AFQT), a commonly used measure of cognitive ability. The particular sample used in the estimations is the nationally representative cross-sectional sample of men. Additional sampling restrictions and the definition of variables used in the regressions are summarized in the Appendix.

Table 1: Average AFQT Score by Occupational Category

Occupation Title	Average AFQT Score	Percent Frequency
Professional and Technical Workers	0.79	0.15
Sales Workers	0.43	0.05
Managers and Administrators	0.42	0.13
Farmers and Farm Managers	0.21	0.01
Clerical and Unskilled Workers	0.16	0.07
High Ability Total	*0.51*	*0.41*
Craftsmen and Kindred Workers	-0.22	0.21
Service Workers	-0.29	0.10
Operatives	-0.37	0.11
Transport Equipment Operatives	-0.43	0.06
Private Household Workers	-0.44	0.00
Farm Laborers and Farm Foremen	-0.51	0.01
Laborers	-0.53	0.10
Low Ability Total	*-0.33*	*0.59*

Note.– Data come from NLSY Men. Occupations are grouped in two based on the average standardized AFQT score of a typical worker in that occupation.

Estimation of the speed of learning requires a substantial number of observations for each occupation category and at each experience level. The small sample size in the NLSY prohibits an accurate estimation for finely defined occupation classes. For this reason, I group the occupations into two broad categories based on ability. Table 1 shows the average AFQT score by single digit occupation classifications.[14] Occupations with a less than average AFQT score are referred to as the low-ability category. The average AFQT scores in high- and low- ability groups are 0.51 and -0.33. High-ability occupations contain 41 percent of workers with a total income share of 0.56.

Table 2 compares the educational attainment in the two occupation groups. Average educa-

---

[14]Reported AFQT scores were standardized by age at the time the test was taken.

Table 2: Education and Ability

Occupation Group:	High Ability	Low Ability
Average Years of Education	14.64	12.02
Education-Ability Gradient	1.26	0.81
	(0.06)	*(0.04)*

Note.– Education-ability gradient is obtained by regressing educational attainment on AFQT score controlling for racial background, parental education and the presence of a library card in the household. Standard errors are reported in parentheses.

tional attainment for the low-ability group is 12.02 years, whereas it is 14.64 years in high-ability occupations. This is expected, since ability and education are positively correlated. The table also reports the results from the regression of educational attainment on AFQT, controlling for background variables. For the high-AFQT group, a one standard deviation increase in the AFQT score leads to an increase of 1.26 years in educational attainment, whereas it leads to an increase of 0.81 years for the low-AFQT occupations.

To estimate the speed of learning, AFQT is regressed on a full set of dummy variables for years of education controlling for racial background. The residuals from this regression, corresponding to ε, represent the component of the AFQT that cannot be inferred by education alone. Then wages are regressed on a full set of experience dummies, the AFQT residual and years of education interacted with experience dummies, and indicators for racial background. Recall that the speed of learning is measured by how fast the coefficients on the residual AFQT rise with experience.

Figure 2 displays the estimated coefficients for each occupation group. The dots represent the coefficients on the AFQT residual, and the dashed line traces the estimated coefficients on educational attainment by experience. In both pictures, the coefficient on ability rises quickly with experience for the first 10-15 years. Employer learning is, therefore, relatively fast, hinting at a minor role for signaling theory. Meanwhile the coefficient on education is approximately stable, which is consistent with the hypothesis that employers' predictions of ability based on education are correct on average.

The speed of learning is estimated by non-linear least squares applied to equation 13 using the estimated coefficients on the AFQT residual by experience. The solid lines in Figure 2 show the fitted learning curves. The estimation results are reported in Table 3. The speed of employer learning is measured by the signal-to-noise ratio, which takes on values between 0 and 1 where 0 represents the absence of learning and 1 stands for immediate learning. The estimate of the speed of learning for high-ability occupations is 0.19, significantly lower than the 0.33 estimated for low-ability occupations. This implies that on-the-job employer learning is slower for high-

Figure 2: The Speed of Learning across Occupational Categories

(a) High-Ability Occupations Decision

(b) Low-Ability Occupations
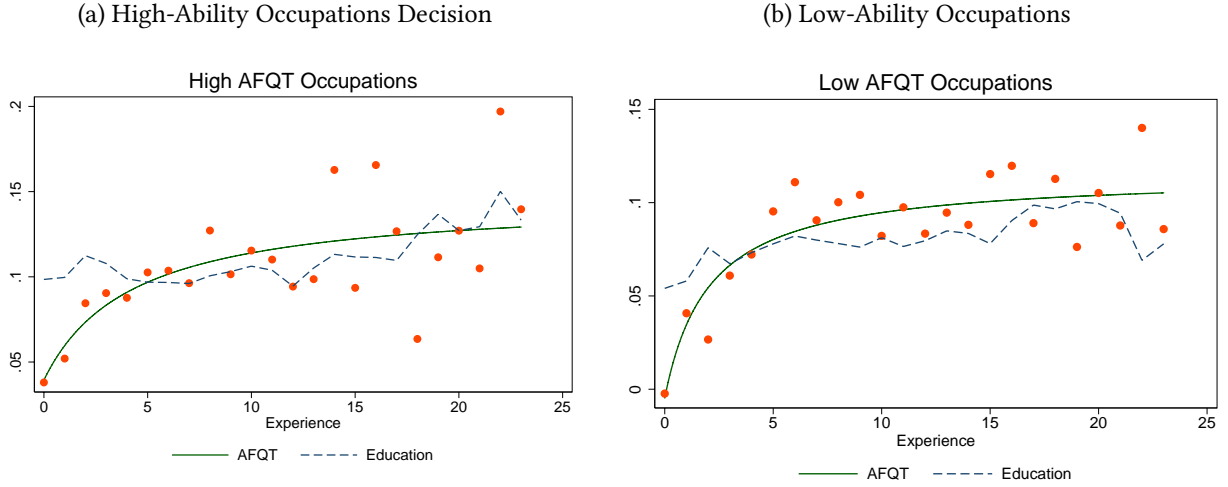
Note.– Figure shows the estimated coefficients from a regression of log wages on residual AFQT score and years of education interacted with indicators for experience. The control variables are indicators for experience and racial background. The solid line shows the fitted learning curve (see Table 3).

ability groups, suggesting a larger potential role for signaling. Notice, however, that the estimated coefficient on the AFQT residual for new entrants, $\kappa_{\bar{z}0}$, is 3.9 percent, indicating that employers in this category also rely on alternative signals to extract information about ability at the time of hire. By contrast, $\hat{\kappa}_{\bar{z}0}$ is essentially zero for the low-ability group, suggesting that prior to employment, employers have little information on the productivity of these workers besides what can be inferred from their education levels. This is not too surprising, since in jobs where it is difficult to assess individual productivity, employers may find it optimal to invest more heavily in recruitment technologies to reduce mistakes in hiring. Tests of aptitude and cognitive ability, often used by firms, are especially useful for this purpose.

Although the coefficients on the AFQT residuals for new entrants are indicative of the extent of recruitment efforts, they likely underestimate the true values of ρ_j in the model for two reasons. First, if employers have access to signals that are correlated with education, but not with cognitive ability, then $\hat{\kappa}_{\bar{z}0j}$ exaggerates the uncertainty faced by employers, and, hence, overestimates the return to signaling. Indirect estimates of ρ_j are, therefore, expected to be higher than those implied by the estimates in Table 3. Second, if workers are equally uncertain about their own abilities, i.e., when the uncertainty is symmetric, then one cannot speak of signaling. Estimating ρ_j allows for this possibility. If learning is symmetric, the estimated speed of learning has no bearing on education decisions. This would be captured in the estimation by a value of ρ_j close to 1.

Table 3: The Speed of Employer Learning

Gradient:	High Ability	Low Ability
$\hat{\kappa}_{\tilde{z}0}$	0.039	-0.005
	(0.024)	(0.016)
$Speed$	0.190	0.328
	(0.151)	(0.105)
$\hat{\kappa}_{\tilde{z}\infty}$	0.146	0.115
	(0.023)	(0.009)

Note.– Results from the non-linear least squares estimation of equation (14). The speed of employer learning takes on values between 0 and 1 where 0 represents the absence of learning and 1 stands for immediate learning. Standard errors are reported in parentheses.

The regression of the standardized AFQT score on indicators for different education levels and race, interacted by indicators for the occupational group yields a correlation coefficient of 0.52. Accordingly, σ_a^2 is set to 0.52.

The change in productivity by experience, captured by $h(x)$, is estimated for each occupational group by regressing log wages on fixed worker effects and a quadratic term in experience. The coefficients on the linear and quadratic components are 5.13 and -0.14 for low-ability occupations and 7.81 and -0.18 for high-ability occupations. The estimated profiles are input directly in the computations.

The remaining parameters of the model are estimated using the simulated method of moments. The procedure is carried out in two steps. First, the differences between the simulated moments and the data moments were minimized for several random draws of ability, a, and occupational taste, τ, for a population of 1,000 workers weighting each moment equally. In the second step, the variance-covariance matrix for parameters is constructed based on the computed simulations, and the first step was repeated weighting the moments by the inverse of the variance-covariance matrix.

4.1   The Return to Job Market Signaling

Table 4 shows the fitted moments. The simulated moments are close to the data moments, but the fit is not perfect due to the non-linearities in the model. The corresponding parameter estimates are reported in the first four columns of Table 5. The results reveal that education better enhances productivity in high-ability occupations. Estimated values for p_1 and p_2 are 0.75 and 0.79, respectively. This suggests that at least part of the differences in educational attainment between the

Table 4: Estimation Results: Moments

	High Ability		Low ability	
	Data	Simulations	Data	Simulations
Mean AFQT	0.51	0.38	-0.33	-0.30
Std. AFQT	0.57	0.64	0.56	0.47
Mean Education	14.64	14.77	12.02	11.92
Education-AFQT Gradient	1.26	1.15	0.81	0.92
Employment Share	0.41	0.44	0.59	0.56

Note.– Results from the SMM estimation. Standard errors are reported in parentheses.

Table 5: Estimation Results: Parameters

A_1	α	r	ρ_1	Λ_1
0.7519	0.0775	0.0806	0.2526	0.2104
(0.0054)	(0.0033)	(0.0016)	(0.0293)	(0.0064)
A_2	β	σ_τ	ρ_2	Λ_2
0.7897	0.3863	0.0078	0.4047	0.2401
(0.0065)	(0.0076)	(0.0001)	(0.0191)	(0.0062)

Note.– Results from the SMM estimation. Standard errors are reported in parentheses.

two occupation groups should be attributed to the human capital role of education.

The estimated effective discount rate is 8.06 percent. This captures not only the typical rate of return on alternative investments, but also other cost factors that are not explicitly modeled here, such as tuition costs. The standard deviation of the occupational preference parameter is estimated to be 7.8 percent of lifetime output per worker, which is comparable to a year of forgone earnings at the margin.

The estimated value for ρ_1 implies that in low-ability occupations, about 25 percent of the uncertainty is unraveled during the recruitment process by interviews or alternative signals. The corresponding estimate for high-ability occupations is 41 percent. These estimates are consistent with the finding that in high-ability occupations, wages of young workers correlate better with test scores conditional on education (Table 3). They thus confirm the earlier conclusion that the recruitment process is more important in occupations that are typically occupied by high-ability workers.

Better recruitment methods in high-ability occupations compensate for the slow on-the-job learning by employers. Combining the estimated values for ρ_1 and ρ_2 with the estimated speeds of employer learning in Table 3 gives the estimates for Λ_j reported in the fifth column of Table 5.

Table 6: Decomposing the Private Return to Education

	Average		Low Ability		High Ability	
Return to Job Market Signaling	2.38	*(0.06)*	2.13	*(0.06)*	2.68	*(0.08)*
Return to Human Capital	6.36	*(0.21)*	6.22	*(0.21)*	6.53	*(0.21)*
Private Return to Education	8.74	*(0.15)*	8.35	*(0.15)*	9.21	*(0.16)*

Note.– Table shows the average estimated marginal return to different functions of education.

Λ_j measures the total extent of uncertainty faced by employers regarding worker abilities. The estimates are similar at 0.24 for high-ability occupations and 0.21 for low-ability occupations.

Recall from equation (5) that $\Lambda \in [0, 1]$ also represents the relative weight of the signaling model. The results show that the contribution of the signaling model to schooling decisions is about 22.4 percent. The remaining 77.6 percent is explained by the human capital model, which is thus the dominant model in the data.

Given the estimates of the production function, and the educational choices, an additional year of schooling raises worker productivity by 6.4 percent on the margin.[15] The marginal return to signaling is 2.7 percent on average for a year of education for workers in high-ability occupations and 2.1 percent for those in low-ability occupations. Given the shares of workers in each group, the average for the entire workforce is 2.4 percent. The average private return to a year of education is, then, 8.8 percent (=2.4 percent + 6.4 percent). Therefore, on the margin, the return to signaling represents 27 percent of the total private return to education. This percentage differs from Λ_j above since it also factors in the relative returns to education and ability. If ability mattered little for a worker's productivity, for instance, the return to signaling would also have been little, regardless of Λ.

The human capital component of the return is in principle comparable to the estimates in the literature that purge the standard OLS estimates of the ability bias. Although a formal analysis is not conducted here, equation (8) hints that an estimate that successfully eliminates ability bias also eliminates the signaling bias, since the latter is contained in the former. Therefore, such an estimate should be compared with the estimated social return to schooling in particular. In this regard, 6.4 percent is at the lower end of the reported estimates: similar to the ones reported by Kaymak (2009) and Angrist and Krueger (1991), but smaller than those reported by Card (1995) (see Card (2001) for a survey).

Under more restrictive assumptions, the human capital role of education can also be tested using wage data on experienced workers, for whom employer learning is mostly completed. At

---

[15]There's a small variation in the marginal return to education arising from finite career length. The reported figure is the average marginal return.

the limit, when the uncertainty about worker productivity is completely unraveled, wages should reflect the true returns to education and ability. Generally, the consistent estimation of these returns requires that the econometrician has access to all productivity signals that the employers have or that these signals are not correlated with educational attainment.[16] This should be kept in mind when interpreting the results below.

The estimated speeds of learning in Table 3 suggest that it takes about 22 years of experience for the uncertainty to be unraveled by 90 percent on average. Using data on workers with more than 22 years of experience, wages are regressed on education controlling for AFQT score, indicators for potential experience, survey year and occupational group (as defined in Table 1.).[17] The estimated return to education is 6.8 percent (1.0 percent), which may still contain a slight upward bias since 10 percent of the uncertainty still remains. Nonetheless, the difference between this estimate and the 6.4 percent estimated above is statistically insignificant, reinforcing the results from the SMM. When split, the estimate for low- and high-ability occupations are 6.1 percent (1.19 percent) and 8.5 percent (2.1 percent). While this is consistent with the finding that education is more productive in high-ability occupations, the difference between the two estimates is slightly higher than what is obtained with the SMM, albeit not statistically significant.

4.2 The Efficiency Cost of Job Market Signaling

The incentive to signal one's ability compels individuals to invest more in education than they would if there were no asymmetric information. The additional investment is excessive since the value of the additional skills acquired does not make up for the cost of schooling. The significance of this externality depends on a multitude of factors, such as the human capital role of education, the extent of asymmetric information, and the cost of schooling.

To gauge the size of the social loss in the model associated with signaling, I simulate two hypothetical economies where the relative weight of signaling, Λ_j, is reduced by 50 percent and by 100 percent. Then using the estimated values of the parameters, I compute counterfactual schooling decisions, occupation choices, and associated productivities for each worker. When the role of signaling is cut by a half, educational attainment declines by 2.6 years on average from 13.1 years to 10.5 years. Given the estimates, this results in a loss of 17 log-points of productivity per year of work. However, lower educational attainment releases more time for work, allowing

---

[16]For instance, a letter of recommendation from one's college professor that vouches for an applicant's attention to detail, a productive trait, is a positive signal of productivity that is also correlated with overall educational attainment. If this additional signal is omitted from the regression, then the estimated return to education with data on experienced workers would still overestimate the human capital role of education. (See Altonji and Pierret (2001) for a detailed treatment of this case.)

[17]Since the NLSY 1979 contains relatively younger workers, the number of observations diminish quickly with experience, leading to higher standard errors. Using workers with more than 25 years of experience yields an estimate of 6.1 (1.6) percent, and using those with more than 27 years of experience yields an estimate of 6.8 (2.8) percent.

each worker to enjoy a longer career. Additional output due to increased work time is 1.2 log-points, resulting in a net loss of 15.8 log-points in lifetime earnings. On the other hand, lower educational attainment induces cost savings. Given the discount rate in the model, the associated savings are 21.2 log-points, resulting in a gain of approximately 5.4 percent. When the asymmetric information is eliminated completely, by setting $\Lambda_j = 0$, the associated net loss in average lifetime output is approximately 7.6 percent.

4.3 Occupational Mobility

Although the model endogenizes the occupational choice at entry, it abstracts from occupational mobility after workers enter the labor market. Allowing occupational mobility may affect the findings here if workers make their education decisions based on the human capital and signaling roles of education along a mobile career path. A related complication arises if workers are not certain of their desired occupation and make their decisions based on an *expected* career path.

A detailed analysis of occupational mobility with employer learning and the associated signaling strategies is beyond the scope of this paper. Nonetheless, the concerns above can in part be addressed using data from the attitudes survey in the NLSY. In 1979, when the survey was initiated, workers were asked what type of job they saw themselves doing when they reached the age of 35. The answers were classified by three-digit occupation categories. Using these answers, it is possible to construct the relevant moments used in the estimation based on *expected* occupation rather than actual occupation. The joint distribution of education and ability by expected occupation provides the appropriate statistics if the NLSY workers decided on their educational attainment according to the type of job they expected to do.

The correlation between expected occupational group and the actual occupation group at the age of 35 is 0.35. The expectations appear somewhat optimistic. Of the workers who expected to work in a low-ability occupation, 75 percent did so and 25 percent worked in a high-ability occupation. Of the workers who expected to work in a high-ability occupation, only 61 percent successfully did, and 39 percent worked in a low-ability occupation.

The results from an SMM estimation using moments by expected occupation are similar to those obtained by using actual occupation. The details of the moments and the corresponding parameter estimates are shown in Tables B.1 and B.2 in the Appendix. Table 7 shows the relative roles of signaling and human capital by occupation. The estimates imply that the average return to signaling is 2.2 percent, which is similar to the 2.4 percent obtained in the benchmark estimation. The rate of increase in human capital on the margin is 7.7 percent, slightly higher than the 6.4 percent obtained in the previous section. That the results are robust to using expected occupation categories instead of actual occupational choices suggests that occupational mobility may not be a

Table 7: Return to Education with Expected Occupation

	Average		Low Ability		High Ability	
Return to Job Market Signaling	2.24	*(0.10)*	2.03	*(0.11)*	2.47	*(0.11)*
Return to Human Capital	7.66	*(0.34)*	7.56	*(0.34)*	7.78	*(0.34)*
Private Return to Education	9.90	*(0.23)*	9.59	*(0.15)*	10.25	*(0.28)*

Note.– Table shows the average estimated marginal return to different functions of education. Results based on expected occupation as reported by the respondent in 1979. See text for details.

concern for the measurement of signaling. Nonetheless, it possible that the reported expectations are only partially responsible for actual educational choices, and a more general model with optimal occupational mobility may yield different results. Understanding the effect of learning on the market's ability to re-assign occupations is a promising venue for future research.

5 Conclusion

The findings suggest that the relationship between educational attainment and wages is better understood through a human capital model, where schooling raises a worker's market productivity. The decisions regarding educational and occupational choices are driven mostly by differences in how valuable the skills acquired in school are in different occupations. The role of the theory of job market signaling in shaping these decisions is non-trivial, albeit quantitatively limited.

This result is driven essentially by two factors. First, in the data, occupations with higher educational attainment, conditional on ability, attract workers with higher ability on average. This type of sorting behavior cannot be reconciled with a labor market described mainly by a signaling model, where one would expect high-ability workers to avoid such occupations. Second, employers appear to learn worker types rather fast, preventing a major role for signaling strategies. Furthermore, where the learning process is slow, employers seem to have developed alternative means of extracting information about worker types, particularly through better recruitment techniques.

References

Altonji, Joseph G. (2005). Employer learning, statistical discrimination and occupational attainment. *American Economic Review*, 95(2), pp. 112–117. doi:10.1257/000282805774670455.

Altonji, Joseph G. and Charles R. Pierret (1997). Employer learning and the signaling value of education. In Isao Ohashi and Toshiaki Tachibanaki, editors, *Internal Labor Markets, Incentives and Employment*. Macmillan Press Ltd., London. doi:10.1057/9780230377974_8.

Altonji, Joseph G. and Charles R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1), pp. 313–350. doi:10.1162/003355301556329.

Angrist, Joshua D. and Alan B. Krueger (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, 106, pp. 979–1014. doi:10.2307/2937954.

Bedard, K. (2001). Human capital versus signaling models: University access and high school dropouts. *Journal of Political Economy*, 109(4), pp. 749–775. doi:10.1086/322089.

Card, David (1995). Using geographic variation in college proximity to estimate the return to schooling. pp. 201–222. Toronto: University of Toronto Press.

Card, David (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5), pp. 1127–1160. doi:10.1111/1468-0262.00237.

Fang, H. (2006). Disentangling the college wage premium: Estimating a model with endogenous education choices. *International Economic Review*, 47(4), pp. 1151–1185. doi:10.1111/j.1468-2354.2006.00409.x.

Farber, Henry S. and Robert Gibbons (1996). Learning and wage dynamics. *The Quarterly Journal of Economics*, 111(4), pp. 1007–1047. doi:10.2307/2946706.

Gibbons, R., L. Katz, T. Lemieux, and D. Parent (2005). Comparative advantage, learning, and sectoral wage determination. *Journal of Labor Economics*, 23(4), pp. 681–724. doi:10.1086/491606.

Kaymak, Barış (2009). Ability bias and the rising education premium in the United States: A cohort-based analysis. *Journal of Human Capital*, 3(3), pp. 224–267. doi:10.1086/651167.

Lang, K. (1994). Does the human-capital / educational-sorting debate matter for development policy? *American Economic Review*, 84(1), pp. 353–358.

Lang, K. and D. Kropp (1986). Human capital vs. Sorting: The effects of compulsory attendance laws. *The Quarterly Journal of Economics*, 101, pp. 609–624. doi:10.2307/1885699.

Lange, Fabian (2007). The speed of employer learning. *Journal of Labor Economics*, 25(1). doi:10.1086/508730.

Riley, John G. (1979a). Informational equilibrium. *Econometrica*, 47(2), pp. 331–360. doi:10.2307/1914187.

Riley, John G. (1979b). Testing the educational screening hypothesis. *Journal of Political Economy*, 87(5), pp. 227–252. doi:10.1086/260830.

Spence, A. Michael (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), pp. 355–379. doi:10.2307/1882010.

Wolpin, Kenneth I. (1977). Education and screening. *American Economic Review*, 67(5), pp. 949–958.

A Data

The data come from the 1979-2004 waves of the National Longitudinal Survey of Youth (NLSY). The analysis here is restricted to the nationally representative cross-sectional sample of 3003 men of all races. The 192 subjects who did not have AFQT scores in the data were dropped. The remaining AFQT scores are standardized within age groups.

The wage variable used in the regressions is the real average hourly rate of pay for the subject's current or most recent job. The consumer price index was used to express wages in 2002 prices. Hourly wage observations less than a $1 and more than $100 were dropped. This resulted in a loss of 16 subjects. The analysis was restricted to jobs after the subject left the school for the first time. This was determined by the first interview when the respondent was not enrolled in school. Invalid observations for educational attainment and observations with less than 8 years of education were dropped. These restrictions resulted in a loss of 206 subjects who were either not finished with their education during the sample or did not have a valid education higher than 8 years. The final data available for regressions consisted of 2588 respondents and 37,136 observations. Further restrictions imposed by the availability of the variables in particular regressions, such as parents' education, are mentioned in the text and the corresponding tables.

The reported estimation statistics are unweighted statistics.

B Additional Tables

Table B.1: Estimation Results with Expected Occupation: Moments

	High Ability		Low ability	
	Data	Simulations	Data	Simulations
Mean AFQT	0.28	0.26	-0.44	-0.21
Std. AFQT	0.58	0.69	0.45	0.67
Mean Education	13.64	13.65	11.83	11.69
Education-AFQT Gradient	1.34	1.16	0.77	0.96
Employment Share	0.61	0.46	0.39	0.54

Note.– Results from the SMM estimation based on expected occupation as reported by respondents in 1979. See text for details.

Table B.2: Estimation Results with Expected Occupation: Parameters

A_1	α	r	ρ_1	Λ_1
0.7318	*0.0726*	0.0941	0.3452	0.1976
(0.0090)	*(0.0024)*	*(0.0025)*	*(0.0436)*	*(0.0103)*
A_2	β	σ_τ	ρ_2	Λ_2
0.7531	*0.4275*	0.0077	0.4742	*0.2240*
(0.0093)	*(0.0118)*	*(0.0002)*	*(0.0257)*	*(0.0089)*

Note.– Results from the SMM estimation. Standard errors are reported in italics. Results based on expected occupation as reported by respondents in 1979. See text for details.