The Speed of Employer Learning

Fabian Lange

Journal of Labor Economics Vol. 25, No. 1 (January 2007), pp. 1–35

James J. Heckman

THE UNIVERSITY OF CHICAGO

Econ 350, Spring 2022

I. Introduction

II. The Empirical Employer-Learning Literature

A. A Test for Statistical Discrimination and Employer Learning

B. Empirical Findings in Altonji and Pierret (2001)

Altonji and Pierret (2001) estimate a log earnings equation that allows for a linear interaction between years of schooling and the AFQT with experience:

$$w_i = \beta_o + \beta_s s_i + \beta_z z_i + \beta_{s,x}(s_i \times x_i) + \beta_{z,x}(z_i \times x_i) + f(x_i) + \beta'_\Phi \Phi_i + \varepsilon_i. \quad (1)$$

- Log wages w_i of individual i depend on schooling s_i
- The AFQT (standardized by birth cohort) z_i
- Experience x_i
- Controls Φ_i

Table 1
The Effects of AFQT and Schooling in a Linear Specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model:								
Education	.0586**	.0829**	.0678**	.0824**	.0887**	.1024**	.0731**	.0846**
	(.0118)	(.0150)	(.0059)	(.0061)	(.0034)	(.0041)	(.0038)	(.0039)
Black	−.1565**	−.1553**					−.0434**	−.0427**
	(.0256)	(.0256)					(.0152)	(.0152)
Female							−.2346**	−.2342**
							(.0092)	(.0092)
Standardized AFQT	.0834**	−.0060	.1010**	.0490**	.1303**	.0686**	.1124**	.0618**
	(.0144)	(.0360)	(.0102)	(.0121)	(.0043)	(.0092)	(.0068)	(.0081)
Education × experience/10	−.0032	−.0234*	−.0030	−.0219**	−.0147**	−.0311**	−.0027	−.0165**
	(.0094)	(.0123)	(.0051)	(.0059)	(.0035)	(.0044)	(.0034)	(.0037)
AFQT × experience/10		.0752**		.0740**		.0729**		.0610**
		(.0286)		(.0119)		(.0099)		(.0077)
R^2	.2861	.2870	.2557	.2588	.1528	.1538	.2988	.3004
Sample	Male, nonhispanic, year < 1993, main and supplementary NLSY sample		Male, white, year < 2000, main NLSY sample		Male, white, year < 2000, main NLSY sample, median regression including zeros		Both genders, year < 2000, main NLSY sample	
No. of individuals	2,978		2,277		2,290		5,336	
No. of observations	21,058		24,410		25,778		55,181	

NOTE.—The coefficients of regressions of log wages on schooling and Armed Forces Qualification Test (AFQT) scores, linearly interacted with the experience coefficient as well as demographic controls, are shown. Columns 1 and 2 report the results reported by Altonji and Pierret (2001) that motivate this story. The specification in Altonji and Pierret (2001) includes a cubic in experience. All specifications examined in this article allow for a full set of experience dummies. Columns 3 and 4 show that the results are found for the sample of white males from the main (nationally representative) sample of the NLSY for the period 1979–98. Columns 5 and 6 investigate whether the results are robust to reinserting the zeros into the sample and performing a median regression. In cols. 5 and 6, I report pseudo-R^2's. Columns 7 and 8 refer to the results obtained on the full sample for the time period 1979–98. For a description of the data, see the appendix. In cols. 1–4, 7, and 8, the standard errors (in parentheses) are White/Huber standard errors accounting for potential correlation at the individual level.

* Statistical significance at the 95% level.

** Statistical significance at the 99% level.

C. Is the AFQT Unobserved by Employers?

III. The Speed of Employer Learning

A. The Employer-Learning Model

- The model specifies individual i's log productivity $x_{i,x}$ at experience x to consist of (i) a linear function $\tilde{x}(s_i, q_i, \eta_i, z_i)$ of various variables (s, q, η, z) describing the information available to employers and researchers and (ii) a polynomial $\widetilde{H}(x_i)$ in experience x_i:

$$\chi_{i,x} = \tilde{\chi}(s_i, q_i, \eta_i, z_i) + \tilde{H}(x_i). \quad (2)$$

The variables s_i capture the information available to both employers and researchers. Schooling is an example for such a variable.

- Log productivity $x_{i,x}$ of individual i at experience level x_i can then be expressed as

$$\chi_{i,x} = rs_i + \alpha_1 q_i + \lambda z_i + \eta_i + \tilde{H}(x_i). \quad (3)$$

The subscript i is understood and will be suppressed from now on.

- Assume that (s, q, η, z) are jointly normally distributed.
- An implication is that the expectation of (z, η) conditional on the information (s, q) available to firms is linear in (s, q)

$$z = E[z|s,q] + v = \gamma_1 q + \gamma_2 s + v; \tag{4}$$

$$\eta = E[\eta|s,q] + e = \alpha_2 s + e. \tag{5}$$

Equations (3)–(5) allow me to express log productivity as a linear function of the information available to employers at time $x = 0$:

$$\begin{aligned} \chi &= (r + \lambda\gamma_2 + \alpha_2)s + (\alpha_1 + \lambda\gamma_1)q + (\lambda v + e) + \tilde{H}(x) \\ &= E[\tilde{\chi}|s,q] + (\lambda v + e) + \tilde{H}(x). \end{aligned} \tag{6}$$

Equations (3)–(5) allow me to express log productivity as a linear function of the information available to employers at time $x = 0$:

$$\chi = (r + \lambda\gamma_2 + \alpha_2)s + (\alpha_1 + \lambda\gamma_1)q + (\lambda\upsilon + e) + \tilde{H}(x)$$

$$= E[\tilde{\chi}|s, q] + (\lambda\upsilon + e) + \tilde{H}(x). \quad (6)$$

The process of employer learning is modeled by assuming that after each period the individual spends in the labor market, a noisy measurement y_τ of $\tilde{\chi}$ becomes available to all employers:

$$y_\tau = \tilde{\chi} + \varepsilon_\tau. \qquad (7)$$

The noise ε_τ is uncorrelated with all other variables in the model.

- At experience x the posterior distribution is normal with mean μ_x and precision $\rho_x = 1/\sigma_x^2$, where $(\mu_x, 1\sigma_x)$ are

$$\mu_x = (1 - \theta_x)E[\tilde{\chi} | s, q] + \theta_x \left(\frac{1}{x} \sum_{\tau=0}^{x-1} y_\tau \right) \tag{8}$$

and

$$\frac{1}{\sigma_x^2} = \frac{1}{\sigma_0^2} + \frac{x}{\sigma_\varepsilon^2}. \tag{9}$$

- The regression coefficients at θ_x each experience level are given by

$$\theta_x = \frac{xK_1}{1 + (x - 1)K_1}, \tag{10}$$

where

$$K_1 = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_\varepsilon^2} \tag{11}$$

is a parameter that reflects the relative information content of initial information (s, q) and subsequent measurements y_τ.

- Therefore, wages equal the expected productivity conditional on the available information:

$$W(s,q,y^x) = E[\exp(\chi)|s,q,y^x]. \tag{12}$$

The distribution of χ conditional on (s,q,y^x) is normal.

- Taking logs and using equation (8) results in the following expression for log wages:

$$w(s,q,y^x) = (1-\theta_x)E[\tilde{\chi} \mid s,q] + \theta_x\left(\frac{1}{x}\sum_{\tau=0}^{x-1} y_\tau\right) + H(x). \qquad (13)$$

B. Estimating the Speed of Learning K1

Without loss of generality we can define the linear projections of (q, η) on (s, z):

$$q = \gamma_3 s + \gamma_4 z + u_1; \tag{14}$$

$$\eta = \gamma_5 s + \gamma_6 z + u_2. \tag{15}$$

- The following equation shows the linear projection of log wages conditional on the observed data *(s, z, x)*:

$$E^*[w(s,q,y^x)|s,z,x] = E^*[(1-\theta_x)E[\tilde{\chi}|s,q] + \theta_x\left(\frac{1}{x}\sum_{\tau=0}^{x-1} y_\tau\right) + H(x)|s,z,x]. \quad (16)$$

The independence assumption on ε_x allows me to write

$$E^*[w(s,q,y^x)|s,z,x] = (1-\theta_x)E^*[E[\tilde{\chi}|s,q]|s,z] + \theta_x E^*[\tilde{\chi}|s,z] + H(x). \quad (17)$$

Equations (6) and (14) imply

$$E^*[E[\tilde{\chi}|s,q]|s,z] = \{\underset{[A]}{r} + \underset{[B]}{\alpha_1\gamma_3} + \underset{[C]}{(\alpha_2 + \lambda\gamma_2 + \lambda\gamma_1\gamma_3)}\}s + \{\underset{[D]}{(\alpha_1 + \lambda\gamma_1)\gamma_4}\}z. \tag{18}$$

- Rewrite $E^*[\tilde{X}|s,z]$ by inserting the linear projections (eqq. [14] and [15]) in equation (3) as

$$E^*[\tilde{\chi}|s,\text{z}] = \{\underset{[E]}{r} + \underset{[F]}{(\alpha_1\gamma_3 + \gamma_5)}\}s + \{\underset{[G]}{\lambda} + \underset{[H]}{(\alpha_1\gamma_4 + \gamma_6)}\}\text{z}. \tag{19}$$

- The two components *[E]* and *[G]* simply reflect the direct productivity effects of schooling and ability, respectively.

Rearranging terms results in

$$E^*[w(s, q, y^x)|s, z, x] = ((1 - \theta_x)b_{s,0} + \theta_x b_{s,\infty})s$$
$$+ ((1 - \theta_x)b_{z,0} + \theta_x b_{z,\infty})z + H(x). \qquad (20)$$

The weights $\theta_x = xK_1/[1 + (x - 1)K_1]$ are functions of K_1 and experience x only.

- Thus, the linear projections (eq. [20]) depend only on K_1 and the following four parameters:

$$b_{s,0} = r + \alpha_1\gamma_3 + \alpha_2 + \lambda(\gamma_2 + \gamma_1\gamma_3); \tag{21a}$$

$$b_{s,\infty} = r + \alpha_1\gamma_3 + \gamma_5; \tag{21b}$$

$$b_{z,0} = (\alpha_1 + \lambda\gamma_1)\gamma_4; \tag{21c}$$

$$b_{z,\infty} = \lambda + \alpha_1\gamma_4 + \gamma_6. \tag{21d}$$

C. Implementation

- The estimating equation corresponding directly to equation (20) regresses log wages on schooling *s* and ability *z* interacted with a complete set of experience dummies:

$$\log\left(w_{t,x}\right) = \Sigma_x \beta_{s,x}(sD_x) + \Sigma_x \beta_{z,x}(zD_x) + \beta'_{\Phi}\Phi_{i,t} + \varepsilon_x. \qquad (22)$$

The controls $\Phi_{i,t}$ include demographic variables and year dummies.

- The parameters $\{\beta_{s,x}, \beta_{z,x}\}_{x=0}^{T}$ are known functions of the structural parameters $\{b_{s,0}, b_{s,\infty}, b_{z,0}, b_{z,\infty}, K_1\}$:

$$\{\beta_{s,x}, \beta_{z,x}\}_{x=0}^{T} = \{(1-\theta_x)b_{s,0} + \theta_x b_{s,\infty} (1-\theta_x)b_{z,0} + \theta_x b_{z,\infty}\}_{x=0}^{T}. \qquad (23)$$

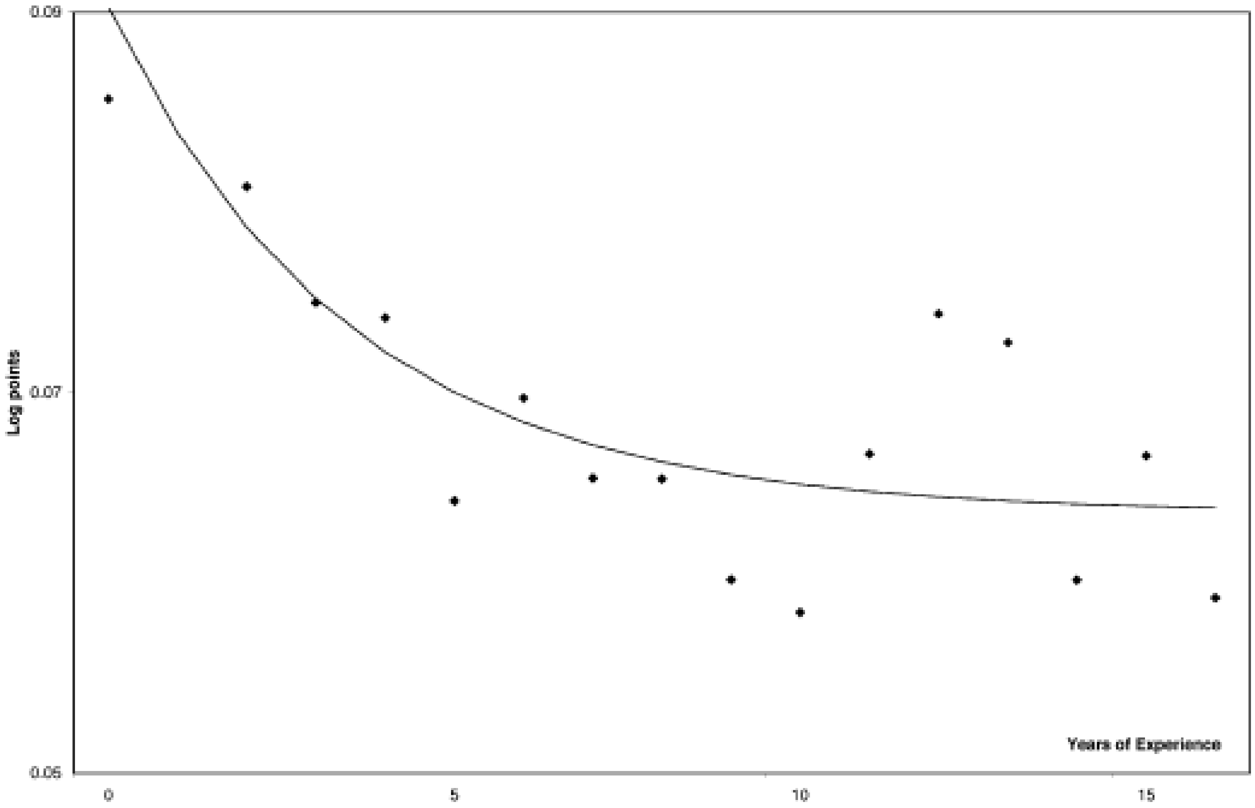

FIG. 1.—Returns to schooling over the life cycle. The scatter displays the estimated coefficients on schooling for each experience level estimated using equation (22). The line shows the predicted returns to schooling over the life cycle implied by the estimates in table 2, column 1. The estimation of these parameters is described in Section III.

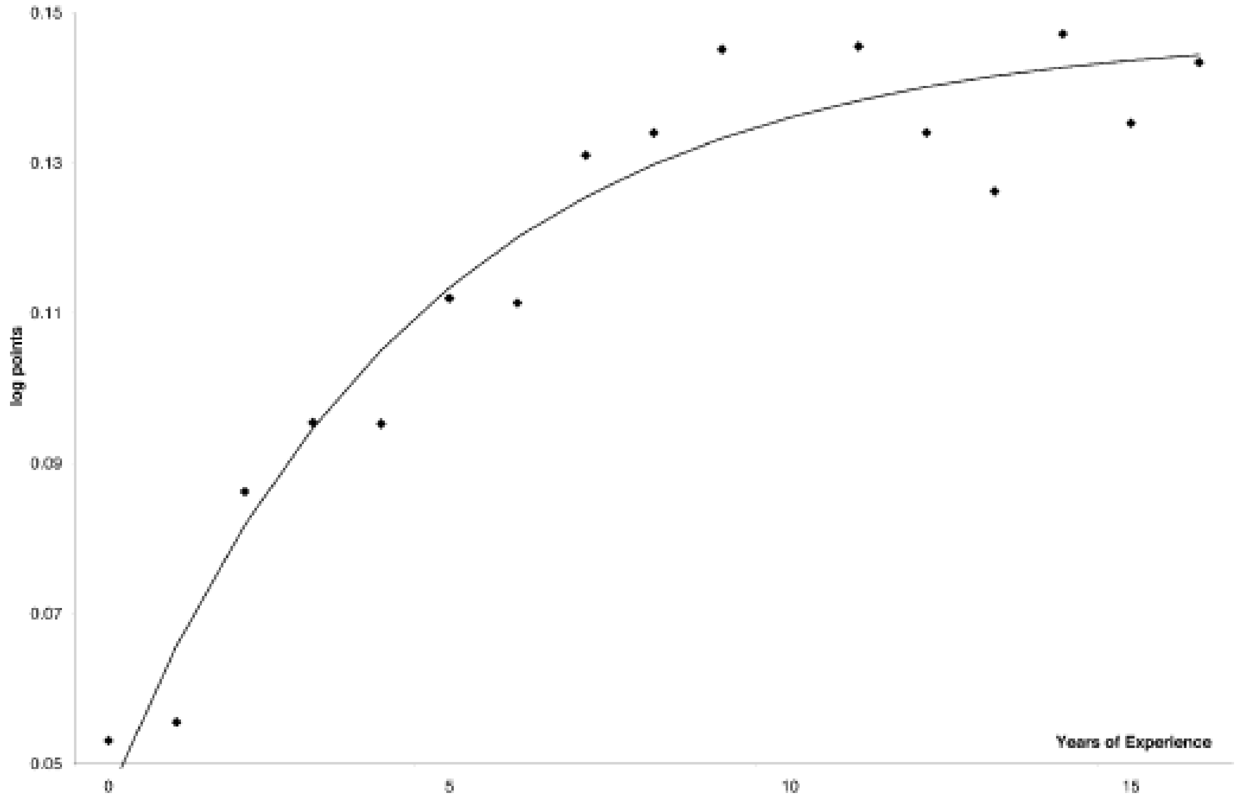

FIG. 2.—Returns to ability over the life cycle. The scatter displays the estimated coefficients on the standardized AFQT score for each experience level estimated using equation (22). The line shows the predicted returns to schooling over the life cycle implied by the estimates in table 2, column 1. The estimation of these parameters is described in Section III.

Table 2
The Speed of Employer Learning

Full Sample (Both Genders, All Races)	(1) Schooling	(2) AFQT Score	(3) Schooling	(3) AFQT Score
Speed of learning K_1	.2891	.2293	.2592	
	(.1139)	(.0860)	(.0922)	
Difference in the estimated K_1's	.0598			
	(.0996)			
Initial value b_0	.1078	−.0044	.1043	−.0104
	(.0152)	(.0303)	(.0107)	(.0329)
Limit value b_∞	.0538	.1772	.0525	.1707
	(.0047)	(.0164)	(.0051)	(.0164)

NOTE.—The reported parameters are estimated by nonlinear least squares using the coefficient estimates on schooling and Armed Forces Qualification Test (AFQT) score at different experience levels obtained from eq. (20) in the text. Section III describes the link between the parameters reported here and the estimated coefficients. The standard errors are obtained by bootstrapping with 5,000 repetitions. Columns 1 and 2 report the results obtained from schooling and the AFQT coefficient separately. Column 3 shows the parameter estimates obtained from using the coefficients on schooling and AFQT jointly.

IV. Job-Market Signaling and the Speed of Learning

A. Defining the Parameter of Interest

Let i denote the rate with which labor earnings are discounted. Then the expected lifetime earnings until retirement T of an individual with characteristics (s, q, z, h) at x p 0 are

$$\int_0^T \exp(-i\tau) E[W(s, q, y^\tau) | s, q, z, \eta] d\tau$$

$$= \int_0^T \exp(-i\tau) E[\exp(E[\tilde{\chi} | s, q, y^\tau] + H(\tau)) | s, q, z, \eta] d\tau. \qquad (24)$$

Differentiating equation (24) with respect to s delivers the increase in the present value of earnings due to an increase in schooling:[11]

$$\int_0^T \exp(-i\tau)\frac{\partial E[W(s,q,y^\tau)|s,q,z,\eta]}{\partial s}d\tau$$

$$= \int_0^T \exp(-i\tau)\frac{\partial E[\exp((1-\theta_\tau)E[\tilde{\chi}|s,q]+\theta_\tau\tilde{\chi}+H(\tau))|s,q,z,\eta]}{\partial s}d\tau \quad (25)$$

$$= \int_0^T \exp(-i\tau)\{E[W(s,q,y^\tau)|s,q,z,\eta]((1-\theta_\tau)(\lambda\gamma_2+\alpha_2))+r\}d\tau.$$

- I therefore define the parameter of interest as

$$F_{\text{JMS}} = \frac{\int_0^T \exp(-i\tau)E[W(s,q,y^\tau)|s](1-\theta_\tau)(\lambda\gamma_2+\alpha_2)d\tau}{\int_0^T \exp(-i\tau)E[W(s,q,y^\tau)|s]((1-\theta_\tau)(\lambda\gamma_2+\alpha_2)+r)d\tau}. \qquad (26)$$

B. Identification Based on the Employer-Learning Model

- The set of equations (21a)–(21d) shows these estimates as functions of the underlying parameters of the model.
- I will restate this set of equations here:

$$b_{s,0} = r + \alpha_1\gamma_3 + (\lambda\gamma_2 + \alpha_2) + \lambda\gamma_1\gamma_3; \quad (21a')$$

$$b_{s,\infty} = r + \alpha_1\gamma_3 + \gamma_5; \quad (21b')$$

$$b_{z,0} = (\alpha_1 + \lambda\gamma_1)\gamma_4; \quad (21c')$$

$$b_{z,\infty} = \lambda + \alpha_1\gamma_4 + \gamma_6. \quad (21d')$$

The residual in log wages is

$$v_{i,x} = w(s, q, y^x) - E[w(s, q, y^x)|s, z, x]. \quad (27)$$

Equation (27) can be expressed as

$$v_{i,x} = (\alpha_1 + (1 - \theta_x)\lambda\gamma_1)u_1 + \theta_x u_2 + \theta \sum_{j=1}^{x} \frac{\varepsilon_j}{x}. \quad (28)$$

Rearranging equations (21a′) and (21b′) yields some insights into the difficulties in identifying $(\lambda\gamma_2 + \alpha_2, r)$:

$$\lambda\gamma_2 + \alpha_2 = (b_{s,0} - b_{s,\infty}) + \gamma_5 - \lambda\gamma_1\gamma_3; \quad (21a'')$$

$$r = b_{Ts} + \alpha_1\gamma_3 - \gamma_5. \quad (21b'')$$

C. Identification Using the Schooling Decision

- Optimal schooling decisions require that the gains from an additional year of schooling must equal the costs from an additional year of schooling.

- Thus, using equation (25),

$$\int_0^T \exp(-i\tau) E[W(s,q,y^\tau)|s,q,z,\eta]((1-\theta_\tau)(\lambda\gamma_2+\alpha_2)+r)d\tau = \phi(s). \quad (29)$$

From equation (25), $\partial w(s,q,0)/\partial s = (\lambda\gamma_2+\alpha_2)+r.$

Therefore, condition $\partial E[w(s,q,0)|s]/\partial s \geq \partial w(s,q,0)/\partial s$ implies

$$\int_0^T \exp(-i\tau) E[W(s,q,y^\tau)|s,q,z,\eta]\left((1-\theta_\tau)\left(\frac{\partial E[w(s,q,0)|s]}{\partial s}-r\right)+r\right)d\tau \geq \phi(s). \quad (30)$$

Averaging equation (29) within schooling levels, then, results in

$$\int_0^T \exp(-i\tau)E[W(s,q,y^\tau)|s]\left((1-\theta_\tau)\left(\frac{\partial E[w(s,q,0)|s]}{\partial s}-r\right)+r\right)d\tau \geq \phi(s). \tag{31}$$

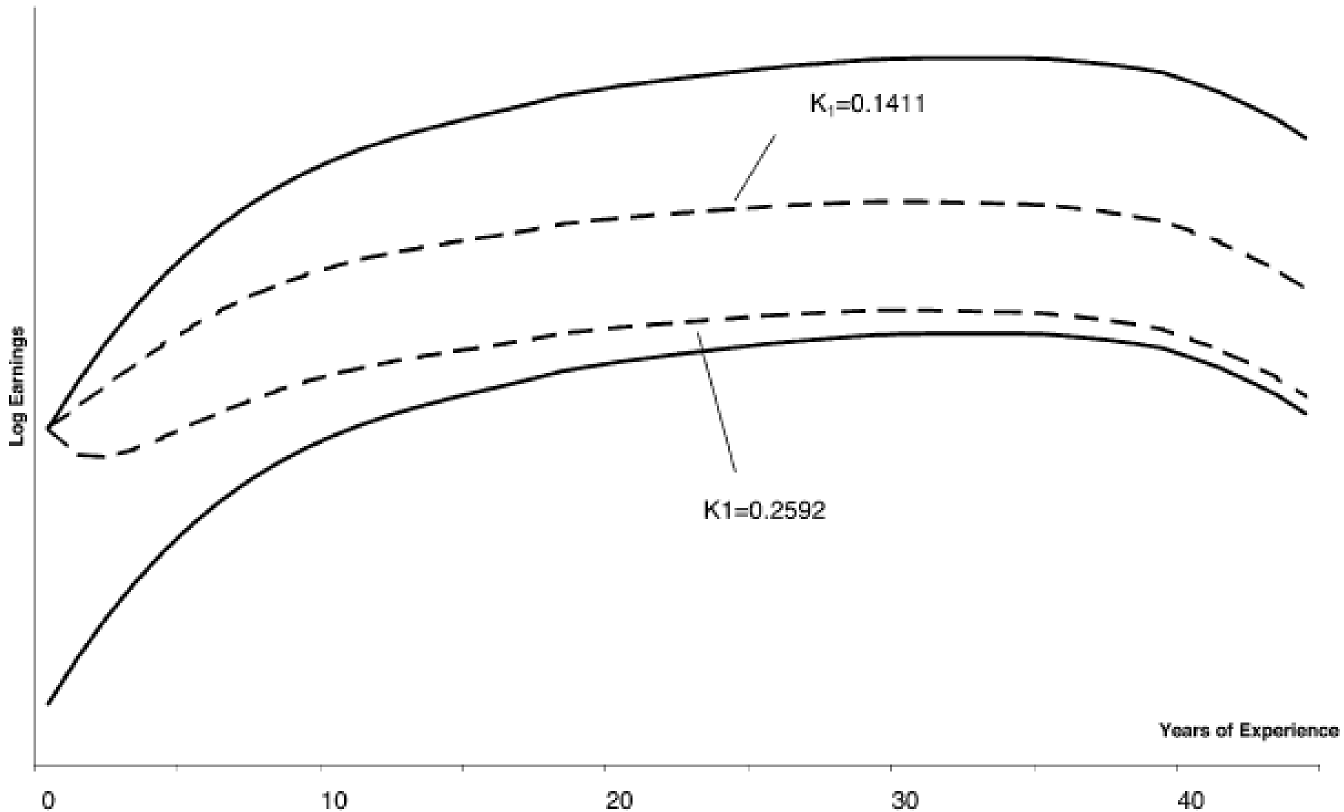

FIG. 3.—Expected earnings in the pure signaling mode. The solid lines show average earnings profiles for high school and college graduates. The dashed lines show expected earnings for workers with the productivity of an average high school graduate who chooses to graduate from college. The figure is drawn under the assumptions that (i) schooling acts as a pure signal and has therefore no productivity-augmenting effects and (ii) employers do not observe any additional information q about individuals' productivity.

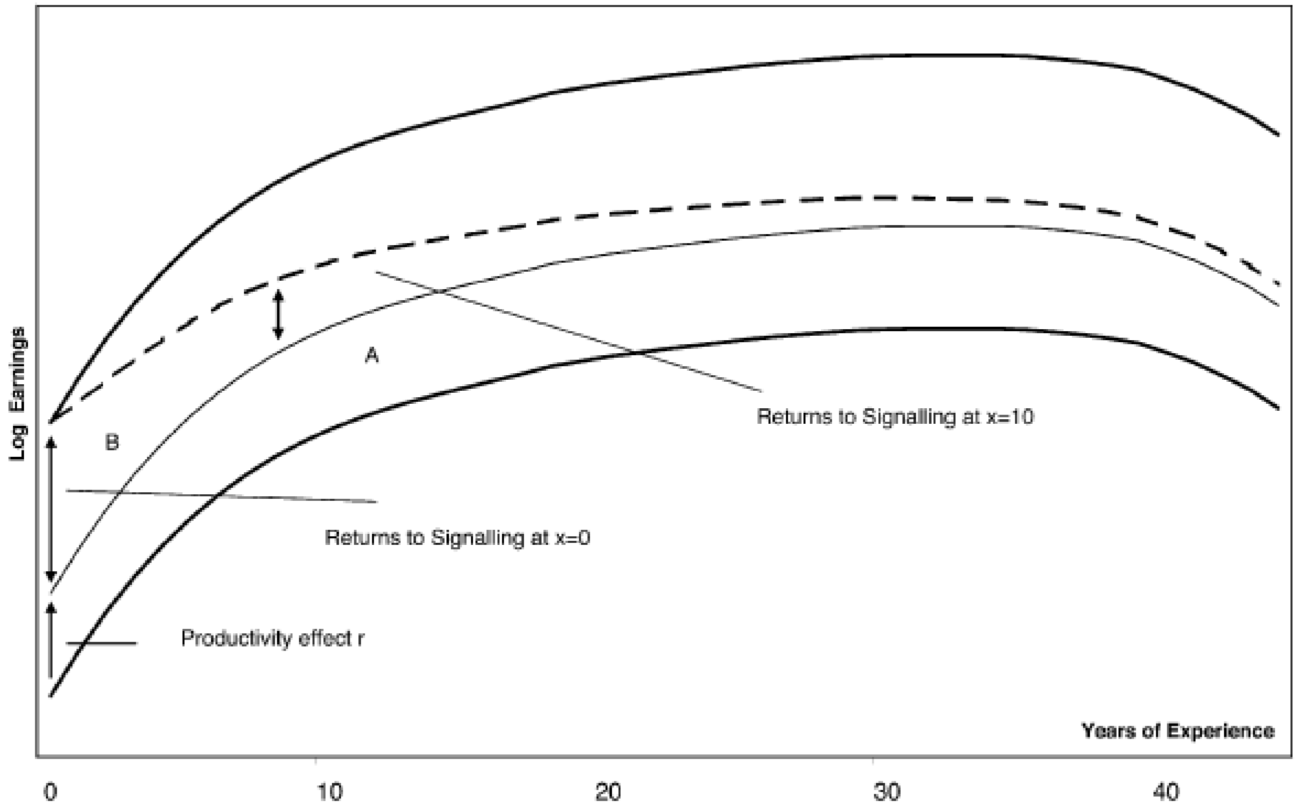

FIG. 4.—Earnings profiles if schooling has signaling and productivity effects. The bold solid lines show average earnings profiles for high school and college graduates. The thin solid line depicts average productivity of high school graduates who decide to attend college. The dashed lines show expected earnings for workers with the productivity of an average high school graduate who chooses to graduate from college. The figure is drawn under the assumption that employers do not observe any additional information q about individuals' productivity.

V. The Bound on the Contribution of Signaling

- The method outlined in the previous section relies on having an estimate of the costs of schooling available.
- I will arrive at this estimate by exploiting the assumption that individuals maximize the present discounted value of lifetime earnings.
- The objective function is then

$$\exp(-is)\int_0^T \exp(-i\tau)E[W(s,q,y^\tau)|s,q,z,\eta]d\tau - \vartheta(s). \qquad (32)$$

Here $\vartheta(s)$ denotes tuition costs of schooling.

- The average marginal costs of schooling among individuals with equal schooling is

$$\phi(s;\, i) = i\int_0^T \exp(-i\tau)E[W(s,q,y^\tau)|s]d\tau + \zeta'(s). \qquad (33)$$

An estimate of $\phi(s;\, i)$ is obtained from observed life-cycle earnings profiles and the 1980 tuition costs of $2,500 reported by Heckman et al. (2003).[14]

- Imposing the inequality (eq. [31]) to hold as an equality and inserting the cost estimate from equation (33) provides the equation from which to solve for the lower bound and the upper bound on the contribution $\underline{r}$ of signaling:

$$\int_0^T \exp(-i\tau) E[W(s, q, y^\tau)|s] \left((1 - \theta_\tau) \left(\frac{\partial E[w(s, q, 0)|s]}{\partial s} - \underline{r} \right) + \underline{r} \right) d\tau = \phi(s; i). \tag{34}$$

Table 3
The Contribution of Signaling to the Gains from Schooling with Estimated Speed of Learning

Interest Rate (%)	Costs/Gains from Schooling (000s)	Contribution of Signaling (%)	Productivity Effects of Schooling (%)
	A. $K_1 = .2592$		
3.00	22.4	25.66	3.4
4.00	22.0	19.57	4.3
5.00	21.5	13.84	5.3
6.00	21.1	8.36	6.5
7.00	20.8	3.10	7.9
8.00	20.4	<.00	8.1
8.70	20.2	<.00	9.3
	B. $K_1 = .1411$		
3.00	22.4	46.77	2.4
4.00	22.0	35.31	3.5
5.00	21.5	24.91	4.7
6.00	21.1	15.03	6.0
7.00	20.8	5.53	7.7
8.00	20.4	<.00	9.5
8.70	20.2	<.00	10.8

NOTE.—Calculations are based on the data for high school graduates. The components needed for this calculation are the speed of learning, the wage profile of high school graduates, the returns to schooling at graduation, and an estimate of tuition costs. The wage profile is estimated from high school wage profiles for experience of 0–18 years and is set constant over the remainder of the life cycle. Individuals are assumed to work for 45 years subsequent to high school graduation (44 if they attend schools an additional year). The wage return at experience = 0 is estimated from a Mincer earnings equation to be 8.70%. Tuition costs are set to $1,900, in line with the numbers reported by Heckman et al. (2003).

VI. Conclusion

Appendix

Table A1
Summary Statistics

Variable	Mean	SD	Min	Max
Highest grade completed	13.12	2.29	6.00	20.00
ln (wage)	6.73	.52	4.61	9.21
Standardized AFQT	.00	1.00	−2.91	2.20
Experience	8.10	4.51	.00	17.00
Hispanic (%)	7.64			
Female (%)	50.11			

NOTE.—AFQT = Armed Forces Qualification Test. Statistics are based on the unweighted cross-sectional sample described in the appendix. The sample consists of 4,701 individuals with 48,930 observations in the years 1979–98.