Employer Learning and Statistical Discrimination

Joseph G. Altonji & Charles R. Pierret. (2001).
Quarterly Journal of Economics, 116(1): 313–350.

Econ 350, Spring 2022

THE UNIVERSITY OF CHICAGO

I. Introduction

II. Implications of Statistical Discrimination and Employer Learning for Wages

II.1 A Model of Employer Learning and Wages

- Our research builds on some previous work, particularly Farber and Gibbons (1996), (hereinafter FG).
- Our model is similar to FG.
- Let y_{it} be the log of labor market productivity of worker i with t_i years of experience:

$$y_{it} = rs_i + \alpha_1 q_i + \Lambda z_i + \eta_i + H(t_i). \quad (1)$$

THE UNIVERSITY OF CHICAGO

- In (1) we separate the determinants of productivity into four categories:
- s_i represents variables that are observed by both the employer and the econometrician;
- q_i includes variables observed by the employer but not seen (or not used) by the econometrician;
- z_i consists of correlates of productivity that are not observed directly by employers but are available to and used by the econometrician;
- and η_i is an index of other determinants of productivity and is not directly observed by the employers and not observed (or observed but not used) by the econometrician.

- Normalize z_i so that all the elements of the conformable coefficient vector Λ are positive.
- In addition, $H(t_i)$ is the experience profile of productivity.
- For now we assume that the experience profile of productivity does not depend on s_i, z_i, q_i, or η_i.

- In the absence of knowledge of z and η, firms form the conditional expectations $E(z|s,q)$ and $E(\eta|s,q)$, which we assume are linear in q and s.
- Consequently,

$$z = E(z|s,q) + v = \gamma_1 q + \gamma_2 s + v \tag{2}$$
$$\eta = E(\eta|s,q) + e = \alpha_2 s + e,$$

- Vector v and the scalar e have mean 0 and are uncorrelated with q and s by definition of an expectation.
- Links from s to z and η may be due in part to a causal effect of s.

- Equations (1) and (2) imply that $\Lambda\nu + e$ is the error in the employer's belief about the log of productivity of the worker at the time the worker enters the labor market.
- The sum $\Lambda\nu + e$ is uncorrelated with q and s.

- Firms do not see y_t, but each period that a worker is in the labor market, firms observe a noisy signal of the productivity of the worker, $\xi_t = y + \epsilon_t$, where $y = y_t - H(t)$.
- ϵ_t reflects transitory variation in the performance of worker i and the effects of variation in the firm environment that are hard for the firm to control for in evaluating the worker.
- Employers know q and s.

- Observing ξ_t is equivalent to observing $d_t = \xi_t - E(y|s,q) = \Lambda\nu + e + \epsilon_t$ which is the sum of the noise ϵ_t and the error $\Lambda\nu + e$ in the employer's belief about initial log productivity.
- The vector $D_t = \{d_l, d_2, \ldots, d_t\}$ summarizes the worker's performance history.
- Let μ_t be the difference between $\Lambda\nu + e$ and $E(\Lambda\nu + e|D_t)$.
- μ_t is uncorrelated with D_t, q, and s.
- μ_t is distributed independently of D_t, q, and s.
- q, s, and D_t are known to all employers, as in FG.

THE UNIVERSITY OF CHICAGO

- As a result of competition among firms, the worker receives a wage W_t equal to $E(Y_t|s,q,D_t)\exp^{\xi_t}$, where Y_t is the level of productivity $\exp^{y_t}$, $E(Y_t|s,q,D_t)$ is expected productivity conditional on s, q, and D_t and $\exp^{\xi_t}$ reflects measurement error and firm-specific factors that are outside the model and are unrelated to s, z, and q.

- Substituting and taking logs, we arrive at the log wage process:

$$w_t = (r + \Lambda\gamma_2 + \alpha_2)s + H^*(t) + (\alpha_1 + \Lambda\gamma_1)q + E(\Lambda v + e|D_t) + \zeta_t, \quad (3)$$

- $w_t = \log(W_t)$ and $H^*(t) = H(t) + \log(E(\exp^{\mu_t}))$.
- $E(\Lambda\nu + e|D_t)$ in (3) shows that wages change over time not just because productivity changes with experience, but also because firms learn about errors in their initial assessment of worker productivity.

- Examine the parameters of the conditional expectation of w_t given s, z, t, and the experience profile $H^*(t)$.
- Begin with the case in which z and s are scalars and then turn to the more general cases.
- Consider the conditional expectation function when $t = 0, \ldots, T$, with

$$E(w_t|s, z, t) = b_{st}s + b_{zt}z + H^*(t). \tag{4}$$

- To simplify the algebra but without any additional assumptions, we reinterpret s, z, and q as the components of s, z, and q that are orthogonal to $H^*(t)$.
- Given that the wage evolves according to (3), the omitted bias formula for least squares regression implies that

$$\begin{aligned} b_{st} = b_{s0} + \Phi_{st} = [r + \Lambda\gamma_2 + \alpha_2] + \Phi_{qs} + \Phi_{st} \\ b_{zt} = b_{z0} + \Phi_{zt} = \Phi_{qz} + \Phi_{zt}, \end{aligned} \tag{5}$$

- where Φ_{qs} and Φ_{qz} denote the coefficients of the auxiliary regressions of $(\alpha_1 + \Lambda\gamma_1)q$ on s and z, respectively, and Φ_{st} and Φ_{zt} are the coefficients of the regression of $E(\Lambda v + e|D_t)$ on s and z.

- Using the facts that $\text{cov}(s, E(\Lambda v + e|D_t)) = 0$ and $\text{cov}(z, E(\Lambda v + e|D_t)) = \text{cov}(v, E(\Lambda v + e|D_t))$ and the least squares regression formula, one may express Φ_{st} and Φ_{zt} as

$$\begin{aligned} \Phi_{st} &= \theta_t \Phi_s \\ \Phi_{zt} &= \theta_t \Phi_z , \end{aligned} \tag{6}$$

- where Φ_s and Φ_z are the coefficients of the regression of $\Lambda v + e$ on s and z and

$$\theta_t = \frac{\text{cov}(E(\Lambda v + e|D_t), z)}{\text{cov}(\Lambda v + e, z)} = \frac{\text{cov}(E(\Lambda v + e|D_t), v)}{\text{cov}(\Lambda v + e, v)}. \tag{7}$$

THE UNIVERSITY OF CHICAGO

Proposition 1. Under the assumptions of the above model,

a. the regression coefficient b_{zt} is nondecreasing in t, and
b. the regression coefficient b_{st} is nonincreasing in t.

Proposition 2. Under the assumptions of the above model,

$$\frac{\partial b_{st}}{\partial t} = -\Phi_{zs} \frac{\partial b_{zt}}{\partial t}.$$

- A matrix version of Proposition 2 still holds

$$\frac{\partial b_{st}}{\partial t} = -\frac{\partial b_{zt}}{\partial t}\Phi_{zs},$$

- where Φ_{zs} is now the $K \times J$ matrix of coefficients of the regression of z on s.

II.2. Statistical Discrimination on the Basis of Race

- If premarket discrimination is an important factor in the gap between the average skills of black and white workers, then it seems likely that various forms of current labor market discrimination contribute to race differences in wages that are unrelated to skill.
- However, it is nevertheless interesting to examine the possibility that a correlation between race and skill might lead a rational, profit-maximizing employer to use race as a cheap source of information about skills.
- Such statistical discrimination along racial lines can have very negative social consequences and is against the law.
- However, it would be hard to detect.
- In contrast, if firms obey the law and do not use race as information, then in the econometric model, race has the properties of a z variable.

- First consider the case where race is the only z variable in the equation.
- In this case our model implies that if (i) race is negatively related to productivity (A ¡ 0), (ii) firms do not statistically discriminate on the basis of race, and (iii) firms learn over time, then (a) the race gap when experience is O will be smaller than if firms illegally use race as information and (b) the race differential will widen as experience accumulates.
- The intuition for (b) is that firms are acquiring additional information about performance that may legitimately be used to differentiate among workers.
- If race is negatively related to productivity, then the new information will lead to a decline in wages.
- If education is negatively related to race, then the coefficient on education should fall with experience.

- We conclude that if firms do not statistically discriminate on the basis of race and race is negatively related to productivity, then (1) the race gap will widen with experience, and (2) adding a favorable z variable to the model will reduce the race difference in the experience profile.

II.3. Alternative Explanations for Variation in the Wage Coefficients with Experience

- The analysis so far assumes that the effects of z and s on the log of productivity do not depend on t.
- Human capital accumulation is included in the model through the $H(t)$ and $H^*(t)$ functions but is assumed to be "neutral" in the sense that it does not influence the experience paths of the effects of s and z on productivity.
- In the more general case, the links between productivity and s and z may depend on experience.
- This would affect the b_{st} and b_{zt}.
- Having a measure of employee training does not by itself allow us to disentangle the effects of learning from those of training.
- To see why, consider the following extension to our basic model.

- The second point is that training may depend on D_t.
- To see the implications of this possibility, suppose that (1) learning is important, (2) variation with s and z in the rate of skill accumulation is not, and (3) variation in our measure of training is driven by worker performance (which leads to promotion into jobs that offer training) rather than by exogenous differences in the level of human capital investment.
- Even under this hypothesis one would expect the introduction of the training measures to lead to a reduction in the growth with t in the coefficient on z and a reduction in the impact of z on the experience path of the coefficient on s.

- For both reasons, we cannot separate the effects of training from the effects of statistical discrimination with learning if, as seems plausible, the quantity of training is influenced by the employer beliefs about productivity.
- With an indicator of Y_t, the identification problem is easily solved, but we lack such an indicator.
- Despite the absence of a clear structural interpretation, we think it is important in this initial study to see how introducing measures of training alters b_{st} and b_{zt}, and we do so below.
- Training may also affect our findings concerning statistical discrimination with respect to race.
- On one hand, ability differences that are correlated with race and that influence the productivity of training may lead the race gap to widen with experience because of differential human capital formation rather than labor market discrimination.

III. Data and Econometric Specification

IV. Results for Education

IV.1. AFQT as a z Variable

Table 1: The Effects of Standardized AFQT and Schooling on Wages

Dependent Variable: Log Wage; OLS estimates (standard errors)

Panel 1 – Experience measure: potential experience				
Model:	(1)	(2)	(3)	(4)
(a) Education	0.0586	0.0829	0.0638	0.0785
	(0.0118)	(0.0150)	(0.0120)	(0.0153)
(b) Black	-0.1565	-0.1553	0.0001	-0.0565
	(0.0256)	(0.0256)	(0.0621)	(0.0723)
(c) Standardized AFQT	0.0834	-0.0060	0.0831	0.0221
	(0.0144)	(0.0360)	(0.0144)	(0.0421)
(d) Education *	-0.0032	-0.0234	-0.0068	-0.0193
experience/10	(0.0094)	(0.0123)	(0.0095)	(0.0127)
(e) Standardized AFQT *		0.0752		0.0515
experience/10		(0.0286)		(0.0343)
(f) Black * experience/l0			-0.1315	-0.0834
			(0.0482)	(0.0581)
R^2	0.2861	0.2870	0.2870	0.2873

Table 1: The Effects of Standardized AFQT and Schooling on Wages

Dependent Variable: Log Wage; OLS estimates (standard errors)

Panel 2 – Experience measure: actual experience instrumented by potential experience				
Model:	(1)	(2)	(3)	(4)
(a) Education	0.0836	0.1218	0.0969	0.1170
	(0.0208)	(0.0243)	(0.0206)	(0.0248)
(b) Black	-0.1310	-0.1306	0.0972	0.0178
	(0.0261)	(0.0260)	(0.0851)	(0.1029)
(c) Standardized AFQT	0.0925	-0.0361	0.0881	0.0062
	(0.0143)	(0.0482)	(0.0143)	(0.0572)
(d) Education *	-0.0539	-0.0952	-0.0665	-0.0889
experience/10	(0.0235)	(0.0276)	(0.0234)	(0.0283)
(e) Standardized AFQT *		0.1407		0.0913
experience/10		(0.0514)		(0.0627)
(f) Black * experience/10			-0.2670	-0.1739
			(0.0968)	(0.1184)
R^2	0.3056	0.3063	0.3061	0.3064

- The key result in the table relating to statistical discrimination is that the coefficient on education $\times t/10$ declines sharply to -.0234 (.0123) when AFQT $\times t/10$ is added between columns (1) and (2).
- The implied effect of an extra year of education declines from .0829 (.0150) to .0595 (.0071) during the first ten years in the labor market.
- These results provide support for the hypothesis that employers have limited information about the productivity of labor force entrants and statistically discriminate on the basis of education.
- Early wages are based on expected productivity conditional on easily observable variables such as education.
- As experience accumulates, wages become more strongly related to variables that are likely to be correlated with productivity but hard for the employer to observe directly.

- While these results give general support for Proposition 1, we may want to know whether the experience profiles of the education and AFQT coefficients satisfy Proposition 2.
- One complication in performing these tests is the place of race within our model-should we treat race as an s variable or a z variable?
- The answer hinges on the extent to which employers violate the law and use race as an indicator of productivity.
- We discuss this at length in Section V below. For now we will sidestep the issue by running separate tests on the white and black samples.

- As noted in Section II, if firms use race as information, then Black behaves as ans variable in the model, and the logic is the same as in our analysis of the effect of education.
- On the other hand, if firms do not use or only partially use race as information, then Black behaves as a z variable.
- In this case the race intercept when experience is 0 will be smaller than when firms use race to discriminate.
- The gap should widen with experience if race is negatively related to productivity, and adding a second z variable that is negatively related to race will: reduce the race gap in experience slopes and possibly make the race intercept more negative.

- The hypothesis that firms do not statistically discriminate on the basis of race does not imply that coefficient on Black will be 0, since race may be correlated with information in q that can legally be used.
- It does imply, however, that the coefficient will be smaller when firms do not use race to discriminate than when they do.
- The fact that the race gap when t equals 0 is essentially 0 and that the gap rises sharply with experience is consistent with the hypothesis of no or very limited statistical discrimination on the basis of race.
- It is inconsistent with the hypothesis that firms make full use of race as information.
- The fact that the coefficient on Black $\times t/10$ rises to -.0834 (.0581) when AFQT $\times t$ is added to the equation (column (4)) is not informative about whether or not firms make full use of race as information.

- The fact that the race gap is so small at low experience levels suggests either that there is not much difference in the productivity of black and white men at the time of labor force entry or that firms do not statistically discriminate very much.
- The accumulation of additional information during a career that can legally be used to differentiate among workers would imply a widening of the race gap with experience (again, if there is a productivity gap) and is fully consistent with our results.
- Another potential test of whether race is used to statistically discriminate or not is to see whether Proposition 2 holds either when race is treated as an s variable or when it is treated as a z variable.
- To do this, we use the model in column (4) of Table III.

V. Do Employers Statistically Discriminate on the Basis of Race?

THE UNIVERSITY OF CHICAGO

VI. Models with Training

Table 2: The Effects of Standardized AFQT, Father's Education, Sibling Wage, Schooling, and Training on Wages

Dependent Variable: Log Wage; Experience Measure: Potential Experience
Training Measure: Predicted before 88, Actual After; OLS estimates (standard errors)

Model:	(1)	(2)	(3)	(4)
(a) Education	0.0606	0.0802	0.0651	0.0746
	(0.0119)	(0.0151)	(0.0121)	(0.0155)
(b) Black	-0.1159	-0.1135	0.0241	-0.0028
	(0.0265)	(0.0267)	(0.0616)	(0.0722)
(c) Standardized AFQT	0.0334	-0.0199	0.0338	0.0102
	(0.0150)	(0.0363)	(0.0150)	(0.0420)
(d) Log of sibling's wage	0.1594	0.0716	0.1611	0.0759
	(0.0213)	(0.0357)	(0.0213)	(0.0356)
(e) Father's education/10	0.0460	0.0211	0.0482	0.0353
	(0.0356)	(0.0974)	(0.0354)	(0.0977)
(f) Education *	-0.0231	-0.0392	-0.0260	-0.0339
experience/10	(0.0095)	(0.0123)	(0.0096)	(0.0128)
(g) Standardized AFQT *		0.0460		0.0207
experience/10		(0.0287)		(0.0339)
(h) Log of sibling's wage *		0.1041		0.1001
experience/10		(0.0402)		(0.0402)
(i) Father's education *		0.0205		0.0084
experience/100		(0.0803)		(0.0805)
(j) Black * experience/10			-0.1180	-0.0945
			(0.0476)	(0.0583)
(k) Training: R_t	-0.1143	-0.1095	-0.1115	-0.1091
	(0.0200)	(0.0199)	(0.0199)	(0.0199)
(1) Cumulative training: Σ	0.1881	0.1830	0.1854	0.1827
R_τ	(0.0139)	(0.0139)	(0.0139)	(0.0139)
R^2	0.3188	0.3199	0.3195	0.3202

Table 3: Estimates of the Effects of AFQT, Father's Education, Sibling Wage, and Schooling on Wage Growth with Controls for Training

Dependent Variable: Δ log Wage; Experience Measure: Potential Experience

Coefficient estimates (standard errors)

Model:	(1)	(2)	(3)	(4)
Education * Δexperience/10	-0.0060	-0.0694	-0.0106	-0.0729
	(0.0833)	(0.0960)	(0.0832)	(0.0959)
AFQT * Δexperience/10		0.3025		0.2975
		(0.1613)		(0.1614)
Log of sibling wage * Δexperience/10		0.2153		0.2107
		(0.1477)		(0.1477)
Father's education * Δexperience/10		-0.4306		-0.4215
		(0.5034)		(0.5034)
Black * Δexperience/10	-0.0504	-0.0425	-0.0503	-0.0426
	(0.0484)	(0.0485)	(0.0483)	(0.0484)
Training: $R_t/10$			0.2468	0.2429
			(0.1024)	(0.1025)
Lag training: $R_{t-1}/10$			-0.0194	-0.0230
			(0.1108)	(0.1108)
S.E.E.	.2965	.2965	.2965	.2964

VII. Conclusions and a Research Agenda