Economics 250a

Learning in the Labor Market

Some References

a) public (symmetric) learning

Henry Farber and Robert Gibbons (1996). "Learning and Wage Dynamics". QJE 1996: 1007-1047

Joseph Altonji and Charles Pierret (2001). "Employer Learning and Statistical Discrimination" QJE: 313-350

Fabian Lange (2007). "The Speed of Employer Learning" JOLE: 1-35

b) asymmetric learning

Robert Gibbons and Lawrence Katz (1991). "Layoffs and Lemons." JOLE": 351-380

Uta Schoenberg (2007) "Testing for Asymmetric Employer Learning" JOLE: 651-691

Jin Li (2011). "Job Mobility, Wage Dispersion, and Asymmetric Information". Unpublished Working Paper, Kellogg School of Management.

c) empirical issues/applications

Thomas Lemieux (2006). "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill? AER: 461-498.

David Card and Ana Rute Cardoso (2012)."Can Compulsory Military Service Raise Civilian Wages? Evidence from the Peacetime Draft in Portugal." AEJ-Applied: 57-93

Introduction

The standard cross-sectional earnings model used by labor economists to address wage-related issues has the form:

$$\log w_i = c + f(Exp_i; d) + \beta S_i + e_i$$

where Exp_i represent $i's$ experience (years of work/years since entering the labor market) and S_i represents $i's$ education (measured in years). This is known as a "human capital earnings function" (HCEF) or "Mincer wage regression". The derivation of this equation will be discussed in Economics 250b. An important fact about HCEF's is that $var[e_i]$ is typically much larger for older workers i.e., the model is conditionally heteroskedastic. An "informal" explanation for this is that

$$e_i = \psi(Exp_i) \cdot a_i + \epsilon_i$$

where a_i is $i's$ "ability", and $\psi(Exp_i)$ is a 'loading factor' that rises with experience, reflecting the fact that as people are in the labor market longer, their "true ability" is revealed to the market, and they are rewarded accordingly. This model implies that

$$var[e_i|Exp_i] = (\psi(Exp_i))^2 \cdot var[a_i] + var[\epsilon_i|Exp_i]$$

which will be rising with experience if $\psi'(.) \geq 0$. (Another explanation for heteroskedasticity is that $var[\epsilon_i]$ rises with experience reflecting the sorting of people to jobs with better person-specific match quality, and the assumption that workers receive some share of match quality rents).

Tables 1a and 1b at the end of the notes are from Lemieux (2006), and show the residual standard deviation of log wages for men and women in different education and experience groups. Notice that this rises a LOT with experience.(It's also true residual standard deviation is much higher for better educated groups). The same pattern is true in longitudinal data, as we will see in various papers to be discussed.

Aside - Statistical Learning Models (Bayesian updating)

In many applications a decision maker is uncertain about the true value of some key parameter, and receives new information over time about the value of the parameter. The natural way to model this class of problems is using Bayesian updating with conjugate priors. The classic reference is de Groot, 1970.

Normal learning.

True state variable is η, with $-\infty < \eta < \infty$. Prior on η is $N(m_0, 1/H_0)$. The observed signal is $s = \eta + \epsilon$, with $\epsilon \sim N(0, 1/h)$, independent of η. It can be shown that posterior for η is

$$N\left(\frac{H_0 m_0 + hs}{H_0 + h}, \frac{1}{H_0 + h}\right).$$

With a sequence of observations $s_t = \eta + \epsilon_t$, with $\epsilon_t \sim N(0, 1/h)$, the posterior after the 1st observation has mean m_1 and precision H_1 given by these formulas. Preceding sequentially, the posterior after the t^{th} observation, conditional on the mean and precision after the $(t-1)^{st}$, is normal with mean and precision

$$\begin{aligned} m_t &= \frac{H_{t-1} m_{t-1} + h s_t}{H_{t-1} + h} = \frac{H_0 m_0 + h \sum_{k=1}^t s_k}{H_0 + th} \\ H_t &= H_{t-1} + h = H_0 + th \end{aligned}$$

Note that

$$m_t \rightarrow \frac{1}{t} \sum_{k=1}^{t} s_k \quad \text{the mean of the signals up to period } t$$

$$H_t \rightarrow th \quad \text{so } \frac{1}{H_t} \rightarrow \frac{1}{t} var[\epsilon_t] \quad \text{the variance of the mean of the signals up to } t$$

This formula has had many applications in labor economics, e.g. models of learning about match quality. Notice that the mean of the posterior evolves like an AR(1) with a rising coefficient on the mean in the previous period, and a falling coefficient on the most recent signal (which is decreasingly informative, given the accumulation of information).

Beta-Bernoulli

Suppose y_t is distributed as a Bernoulli with $P(y_t = 1) = p$. The conjugate prior for p is $Beta(\alpha, \beta)$. For $p \sim Beta(\alpha, \beta)$:

$$
\begin{aligned}
E[p] &= \frac{\alpha}{\alpha+\beta}, \\
var[p] &= \frac{\alpha\beta}{(\alpha+\beta)^2(1+\alpha+\beta)}
\end{aligned}
$$

The density is $f(p) = \frac{\Gamma(\alpha)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$. Note that $Beta(1,1) = U(0,1)$. The posterior for p, given a draw y_1 is $Beta(y_1 + \alpha, 1 - y_1 + \beta)$. Applying this sequentially, the posterior after t realizations with S_t successes is $Beta(S_t + \alpha, t - S_t + \beta)$, implying the posterior mean and variance are

$$
\begin{aligned}
E[p|S_t] &= \frac{S_t+\alpha}{t+\alpha+\beta} \to \frac{1}{t}S_t \\
var[p] &= \frac{(S_t+\alpha)(t-S_t+\beta)}{(t+\alpha+\beta)^2(1+\alpha+\beta+t)} \to \frac{\frac{1}{t}S_t \times (1-\frac{1}{t}S_t)}{t}
\end{aligned}
$$

A nice application of this class of learning models is to problems of the form "waiting for a prize that will arrive with unknown probability p." If the model is formulated so the agent "opts out" of the wait when p is low, then optimal behavior is to wait until n unsuccessful draws, then opt out. In his thesis, Larry Katz applied this idea to the behavior of workers on temporary layoff, who have to decide whether to continue waiting for recall, or start looking for a new job.

Public (symmetric) learning

Faber-Gibbons (FG) lay out the basic "public learning" model. Their notation is as follows: w_{it} = wage of person i in t^{th} year in labor market (so t = experience); y_{it} = output of person i in year t (assumed to be public); η_i = ability of person i (assumed to be fixed but unknown); S_i, X_i = schooling and some other time invariant observed characteristics; Z_i = invariant characteristic that is observed in the market but not seen by econometrician). Their basic model of output is:

$$y_{it} = y_i + \epsilon_{it}$$

where ϵ_{it} is an i.i.d. shock. They also consider a model with a homogeneous experience effect:

$$y_{it} = y_i + h(t) + \epsilon_{it}.$$

Finally, they assume

$$w_{it} = E[y_i|S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-1}] + h(t) \ .$$

In other words, the wage is the market's best estimate of the time-invariant productivity effect y_i given observed stuff.

This simple setup has 3 key predictions. First, consider a regression of the wage in period t on S_i, X_i :

$$w_{it} = \alpha_t + \beta_t S_i + \gamma_t X_i + e_{it}. \tag{1}$$

Let $E^*[v|u]$ denote the linear projection of v on u. From the assumed wage-setting model (ignoring h) :

$$\begin{aligned} E^*[w_{it}|S_i, X_i] &= E^*[E[y_i|S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-1}] \mid S_i, X_i] \\ &= E^*[E^*[E[y_i|S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-1}] \mid S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-1}] \mid S_i, X_i] \\ &= E^*[E^*[y_i|S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-1}] \mid S_i, X_i] \\ &= E^*[y_i|S_i, X_i], \end{aligned}$$

where we are using the law of iterated projections (line 1 to line 2 and line 3 to line 4) and the fact that $E^*[E[u|v]|v] = E^*[u|v]$. From the last line it follows that the projection coefficients do not depend on t. This means that even though employers are gradually learning (and thus have to rely less and less on S_i, X_i) the projection coefficients from models that ONLY condition on S_i, X_i don't change. This same argument works if we include $h(t)$.

A second prediction comes from comparing wages in consecutive periods:

$$\begin{aligned} w_{it} &= E[y_i|S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-1}] \\ w_{it-1} &= E[y_i|S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-2}]. \end{aligned}$$

Thus we can write:

$$w_{it} = w_{it-1} + \zeta_{it}$$

where

$$E[\zeta_{it}|S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-2}] = 0.$$

The innovation in wages from $t-1$ to t has to be independent of all information available at $t-1$ so that means the wage follows a Martingale process (ie a random walk). Notice that we would also expect $var[\varsigma_{it}]$ to be declining in t. Notice too that if

$$w_{it} = E[y_i|S_i, X_i, Z_i, y_{i1}, y_{i2}, ...y_{it-1}] + h(t)$$

then

$$w_{it} = \Delta h(t) + w_{it-1} + \varsigma_{it}$$

so the experience-adjusted wage residual is a random walk.

A third prediction comes from considering a piece of "background information": B_i that is available to us (the outside observers) but NOT to the market. FG suggest AFQT and the presence of a library card in the family home when a kid was a teenager as two examples of B. Consider the residual of B from a regression on S_i, X_i and the first period wage w_{i1} :

$$B_i^* = B_i - E^*[B_i|S_i, X_i, w_{i1}].$$

and period-specific regressions of wages on S_i, X_i, and B_i^* :

$$w_{it} = \alpha_t + \beta_t S_i + \gamma_t X_i + \pi_t B_i^* + e_{it}.$$

Since by construction B_i^* is orthogonal to the other regressors:

$$\begin{aligned}\pi_t &= \frac{cov[w_{it}, B_i^*]}{var[B_i^*]} \\ &= \frac{1}{var[B_i^*]} cov[w_{i1} + \zeta_{i2} + \zeta_{i3} + .. + \zeta_{it}, B_i^*] \\ &= \frac{1}{var[B_i^*]} \sum_{\tau=2}^{t} cov[\zeta_{it}, B_i^*]\end{aligned}$$

Now assuming that $cov[\eta_i, B_i] > 0$, we would expect that the innovations would be positively correlated with B_i^* : i.e., employers are gradually learning whatever is in B. In this case π_t is an increasing function of t.

In their empirical work, FG use longitudinal data from the NLSY to do 3 things. First, they show that the regression coefficient of wages on schooling is roughly constant with age. Second, they show that if you add information about AFQT and library cards, and allow time-varying coefficients, then the coefficients of both variables rise with experience. Third, they fit a covariance model to wage residuals and test a random walk model (similar to the exercise we did in class). A feature of FG's empirical work that other researchers have not used is the analysis of wage levels rather than logs.

Altonji-Pierret (AP)

AP build on FG – particularly on FG's third implication – but focus what happens to the regression coefficients at different levels of experience from a regression of wages on schooling and B_i (rather than B_i^*). This turns out to have some interesting implications. They also build up a "log wage" model, which is arguably more useful for empirical work than a wage levels model. The starting model for log productivity of person i in period t is:

$$y_{it} = rS_i + \delta z_i + \eta_i + h(t_i) \tag{2}$$

where S_i is schooling, z_i is a set of "background factors" (like AFQT) that are not known to employers but are observed by econometricians, η_i is ability, and $h(t_i)$ is a homogeneous experience effect (as in FG). (Note that apart from the experience part, log productivity is a constant). Employers have a set of variables q_i which they observe and we don't, that they use to set wages. Assume their conditional expectations are linear in (S_i, q_i) :

$$z_i = E[z_i|S_i, q_i] + v_i = \gamma_1 q_i + \gamma_2 S_i + v_i \tag{3}$$

$$\eta_i = E[\eta_i|S_i, q_i] + e_i = \alpha_1 q_i + \alpha_2 S_i + e_i \tag{4}$$

In each period employers observe $y_{it} + \epsilon_{it}$, which is equivalent to observing

$$\begin{aligned}d_{it} &\equiv y_{it} - E[y_{it}|S_i, q_i] \\ &= \delta v_i + e_i + \epsilon_{it}\end{aligned}$$

So if there was no ϵ_{it} they could observe $\delta v_i + e_i$ and form a "perfect" prediction for y_{it}. In period t it is assumed that employers set the wage for individual i equal to the expected level of productivity, given the information set

$$D_{it} = \{q_i, S_i, d_{i1}, d_{i2}, ...d_{it}\}$$

(This timing convention is a little unusual – its as if wages are set at the end of the period). Define

$$\begin{aligned} \mu_{it} &= \delta v_i + e_i - E[\delta v_i + e_i | D_{it}] \\ &= y_{it} - E[y_{it}|D_{it}] \end{aligned}$$

Using (2) and (3a, 3b), actual log productivity is:

$$y_{it} = (r + \delta\gamma_2 + \alpha_2)S_i + (\delta\gamma_1 + \alpha_1)q_i + E[\delta v_i + e_i|D_{it}] + \mu_{it} + h(t_i),$$

so the level of productivity is:

$$\begin{aligned} Y_{it} &= \exp(y_{it}) \\ &= \exp((r + \delta\gamma_2 + \alpha_2)S_i + (\delta\gamma_1 + \alpha_1)q_i) \\ &\quad \cdot \exp(E[\delta v_i + e_i|D_{it}]) \cdot \exp(h(t_i)) \cdot \exp(\mu_{it}) \end{aligned}$$

Taking expectations conditional on D_{it} (and noting that only the last term is random, given D_{it}) :

$$\begin{aligned} E[Y_{it}|D_{it}] &= \exp((r + \delta\gamma_2 + \alpha_2)S_i + (\delta\gamma_1 + \alpha_1)q_i) \\ &\quad \cdot \exp(E[\delta v_i + e_i|D_{it}]) \cdot \exp(h(t_i)) \cdot E[\exp(\mu_{it})|D_{it}] \end{aligned}$$

Assuming that the wage is set equal to expected productivity, and taking logs, we get:

$$\begin{aligned} w_{it} &= \log(wage) = \log E[Y_{it}|D_{it}] \\ &= (r + \delta\gamma_2 + \alpha_2)S_i + (\delta\gamma_1 + \alpha_1)q_i + E[\delta v_i + e_i|D_{it}] + h(t_i) + \log E[\exp(\mu_{it})|D_{it}] \end{aligned}$$

The last term is similar to the error component that arises in taking first differences from a log-linearized intertemporal labor supply function, and would disappear if we could interchange log and E operators. Notice that building up from a log productivity model we get a log wage model with a "learning" component $E[\delta v_i + e_i|D_{it}]$.

Now consider what happens if we regress w_{it} on (S_i, z_i) (after partially out any experience effect):

$$w_{it} = b_{st}s_i + b_{zt}z_i + \phi_{it}.$$

The formulas for the coefficients can be written as:

$$\begin{aligned} b_{st} &= (r + \delta\gamma_2 + \alpha_2) + (\delta\gamma_1 + \alpha_1)\Gamma_{q,s|z} + \Phi_{st} = b_{s0} + \Phi_{st} \\ b_{zt} &= (\delta\gamma_1 + \alpha_1)\Gamma_{q,z|s} + \Phi_{zt} = b_{z0} + \Phi_{zt} \end{aligned}$$

where $\Gamma_{q,s|z}$ is the regression coefficient from a regression of q on s, partialling out the effect of z, and $\Gamma_{q,z|s}$ is the parallel regression coefficient from a regression of q on z, partialling out the effect of s, and the Φ terms are coefficients from the auxillary regression:

$$E[\delta v_i + e_i | D_{it}] = \Phi_{st} \cdot s_i + \Phi_{zt} \cdot z_i + \kappa_{it}.$$

A first observation is that $\Phi_{s0} = \Phi_{z0} = 0$ because the information set at time 0, D_{i0}, only has (s_i, q_i), and v_i and e_i are orthogonal to (s_i, q_i) (see equations 3a, 3b). The terms b_{s0} and b_{z0} reflect the projection of what employers see at the beginning of a career on what we (as econometricians) see. A second set of results – the main results derived by AP – are that Φ_{st} and Φ_{zt} have a very special structure:

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

where

$$\theta_t = \frac{cov[E[\delta v_i + e_i | D_{it}], v_i]}{cov[\delta v_i + e_i, v_i]}$$

and Φ_s and Φ_z are the coefficients from an auxilliary regression:

$$\delta v_i + e_i = \Phi_s \cdot s_i + \Phi_z \cdot z_i + \kappa_i'$$

with

$$\Phi_s = -\Phi_z \Gamma_{z,s}$$

Notice that with learning over time, θ_t is rising toward 1. As this happens, the coefficient on schooling is declining while the coefficient on z_i (i.e., AFQT) is rising. Morever, there is a strong testable implication that the change in one of these coefficients is related to the change in the other by the observable factor $\Gamma_{z,s}$ - which is just the coefficient from a univariate regression of z on s. Intuitively, at the start of the career, the observed wage appears to be "too strongly" correlated with the observable factor s and "too weakly" correlated with the unobserved factor z. As employers learn they "unload" some of the explanatory power from s to z.

This special structure all arises because

$$cov[s_i, E[\delta v_i + e_i | D_{it}]] = 0$$

and

$$cov[z_i, E[\delta v_i + e_i | D_{it}]] = cov[v_i, E[\delta v_i + e_i | D_{it}]]$$

It is not hard to see that if you have a set of regression models like:

$$y(k) = a + b_1 x_1 + b_2 x_2 + e$$

and $cov[y(k), x_1] = 0$, then

$$\begin{aligned} b_2(k) &= \frac{cov[y(k), x_2]}{var[x_2|x_1]} \\ &= \frac{cov[y(k), x_2]}{cov[y(0), x_2]} \times \frac{cov[y(0), x_2]}{var[x_2|x_1]} \\ &= \frac{cov[y(k), x_2]}{cov[y(0), x_2]} \times b_2(0) \end{aligned}$$

This is the source of the "θ_t" structure for Φ_{zt}. As an exercise, show that in the case where $cov[y(k), x_1] = 0$, there is a simple relation between $b_2(k)$ and $b_1(k)$.

AP use this setup to address the question of whether employers "statistically discriminate" against young black workers by using race as an "s" variable that is negatively correlated with AFQT (an obvious z variable). If they do so, then we would expect the coefficient on race to fall (in magnitude), and the coefficient on AFQT to rise in magnitude as employers learn about the proportion of true abiltity that is contained in AFQT. They also try a related analysis in which a sibling's wage (or father's education) are treated as "z" variables. Overall they find that the effect of black race increases (in magnitude) with experience, while the effect of the "hidden" variables also tend to rise. (See Tables 1 and 2 from their paper at the end of the notes).

Lange (2006) uses the set up of AP plus the "normal learning model" at the start of the notes to parameterize the rate at which the coefficients on "s" and "z" change with experience. (see his equations 20-21). The implied rate of learning is pretty fast – most of the evolution of the coefficients is completed within 5-7 years of labor market entry.

Asymmetric learning

A question that has long interested labor economists is whether information about employee ability is only directly observable by the firm (or firms) who employ the worker. Such a possibility sets up a potential "winner's curse", in which outside firms can only bid away workers who are worth less than they have been offered. One solution is to assume (as in Gibbons and Katz) that some workers leave for exogenous reasons. Another possibility – assumed in Li (2011) – is that uninformed outsiders bid using randomized strategies. Some of the setup in Li's paper is very close to models of affiliated value auctions in which there is a better informed bidder, as in Hendricks and Porter (AER, 1988).

Gibbons-Katz (GK)

GK have a 2-period model. A worker has productivity $\eta \sim F$, with $\eta_L \leq \eta \leq \eta_H$. Prior to period 1, everyone is uninformed and workers are randomly assigned to firms. In period 1 a worker has output η – so the incumbent employer is perfectly informed. In period 2, the workers output will be:

$$
\begin{array}{ll}
\eta + s & \text{if she stays with incumbent} \\
\eta & \text{if she moves to a new firm}
\end{array}
$$

where $s > 0$ reflects some kind of firm-specific gain. It is assumed that $E[\eta] > \eta_L + s$ (for reasons described below). The timing between period 1 and 2 is as follows:

- the incumbent decides to lay off workers with $\eta \leq \eta_R$
- the non-laid off receive offers from outside bidders. The bid is w_m, which will be a function of η_R
- the incumbent matches the offer for any worker with $\eta + s \geq w_m$, declines to match for the others
- a fraction μ of the non-laid-off leave regardless of their wage offer
- the remaining fraction $1 - \mu$ of the non-laid-off leave iff their offer is not matched.

GK first solve for the optimal offer of outsiders. Assuming competition, this has to satisfy:

$$
\begin{aligned}
0 &= \mu(E[\eta|\eta \geq \eta_R] - w_m) + \\
(1-\mu)\Pr[\eta + s &< w_m | \eta \geq \eta_R] \cdot (E[\eta|\eta_R \leq \eta \leq w_m - s] - w_m)
\end{aligned}
$$

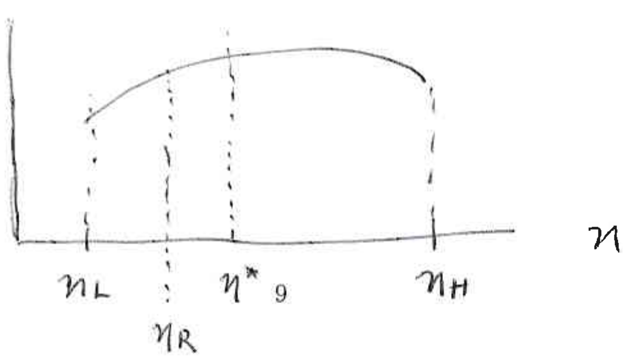
The first term is the profit to be made on exogenous leavers. These are positively selected ($\eta \geq \eta_R$). The second term is the profit (or loss, actually) to be made on the workers who are left unmatched. This equation – GK's equation (1) – has a solution for relevant values of η_R. Notice next that given w_m, there will in fact be a range of $\eta' s$ who are not laid off, but who are "unmatched", iff

$$w_m \geq \eta_R + s$$

Finally, define η^* by:

$$\eta^* + s = E[\eta|\eta \geq \eta^*]$$

If F is log concave, then $\frac{d}{d\eta^*}E[\eta|\eta \geq \eta^*] < 1$, and since the lhs of this equation is less than the rhs for $\eta^* = \eta_L$, and is above the rhs when $\eta^* = \eta_H$ there is a solution. GK show that if $\eta_R < \eta^*$ then the solution for w_m will satisfy the condition that $w_m \geq \eta_R + s$. In fact there is range of equilibria with different values for η_R in the interval $\eta_L < \eta_R < \eta^*$ that will all work.



The last question is: what is the wage for laid off workers? In the equilibrium this will be

$$w^{layoff} = E[\eta|\eta < \eta_R]$$

So to summarize: the lowest-productivity workers are laid off / fired. The next group (in the range from η_R to $w_m - s$) are all "allowed to leave" by not having their offers matched. These workers, plus a fraction μ of everyone above the cutoff who quit for exogenous reasons, all leave voluntarily and earn w_m. The "stayers" also all earn w_m.

In their empirical work, GK compare wage changes from the old job to the new job for 2 kinds of movers: people displaced by plant closings, and people who were laid off for other reasons (the "lemons" of their title). They find that the displaced workers do better than the other types of laid off workers.

Table 1a: Within-group variance of wages by experience-education cell for men, 1973-75 and 2000-02

	Within-group variance			Workforce share		
	1973-75	2000-02	Change	1973-75	2000-02	Change
	(1)	(2)	(3)	(4)	(5)	(6)
A. By education and experience						
Dropout:						
1-10	0.118	0.083	-0.035*	0.065	0.035	-0.030
11-20	0.169	0.130	-0.038*	0.052	0.026	-0.026
21-30	0.170	0.154	-0.017*	0.055	0.025	-0.029
31+	0.180	0.162	-0.019*	0.123	0.028	-0.095
High school graduates:						
1-10	0.130	0.130	0.000	0.137	0.082	-0.055
11-20	0.145	0.181	0.035*	0.094	0.085	-0.009
21-30	0.162	0.196	0.034*	0.069	0.086	0.017
31+	0.188	0.217	0.029*	0.074	0.058	-0.016
Some college:						
1-10	0.143	0.152	0.008	0.076	0.077	0.001
11-20	0.173	0.204	0.031*	0.036	0.075	0.039
21-30	0.216	0.227	0.012	0.025	0.072	0.048
31+	0.245	0.256	0.011	0.020	0.046	0.026
College graduates:						
1-10	0.161	0.224	**0.064***	0.048	0.061	0.014
11-20	0.204	0.276	**0.072***	0.022	0.063	0.041
21-30	0.220	0.310	**0.091***	0.017	0.051	0.034
31+	0.299	0.332	0.033	0.009	0.024	0.015
Post-graduates:						
1-10	0.217	0.316	**0.099***	0.034	0.023	-0.010
11-20	0.324	0.324	0.000	0.023	0.033	0.009
21-30	0.327	0.302	-0.025	0.015	0.033	0.018
31+	0.420	0.369	-0.051	0.006	0.016	0.010
B. Weighted Average (using alternative shares)						
Actual shares	0.173	0.214	0.041			
1973-75 shares	0.173	0.185	0.012			
2000-02 shares	0.191	0.214	0.023			

Table 1b: Within-group variance of wages by experience-education cell for women, 1973-75 and 2000-02

	Within-group variance			Workforce share		
	1973-75	2000-02	Change	1973-75	2000-02	Change
	(1)	(2)	(3)	(4)	(5)	(6)
A. By education and experience						
Dropout:						
1-10	0.099	0.056	-0.043*	0.057	0.026	-0.031
11-20	0.130	0.090	-0.040*	0.039	0.015	-0.024
21-30	0.125	0.106	-0.019*	0.050	0.018	-0.032
31+	0.139	0.123	-0.017*	0.103	0.023	-0.080
High school graduates:						
1-10	0.106	0.108	0.002	0.179	0.070	-0.109
11-20	0.145	0.157	0.011*	0.095	0.072	-0.023
21-30	0.144	0.172	0.028*	0.092	0.086	-0.006
31+	0.162	0.178	0.016*	0.097	0.074	-0.023
Some college:						
1-10	0.118	0.137	0.019*	0.077	0.091	0.014
11-20	0.134	0.198	**0.065***	0.025	0.081	0.057
21-30	0.152	0.209	**0.057***	0.020	0.084	0.064
31+	0.160	0.220	**0.060***	0.020	0.054	0.034
College graduates:						
1-10	0.134	0.179	**0.045***	0.055	0.076	0.020
11-20	0.170	0.260	**0.090***	0.015	0.058	0.043
21-30	0.173	0.262	**0.088***	0.014	0.052	0.038
31+	0.195	0.254	**0.059***	0.010	0.021	0.010
College post-graduates						
1-10	0.154	0.239	**0.085***	0.022	0.026	0.004
11-20	0.238	0.259	0.021	0.012	0.027	0.015
21-30	0.204	0.217	0.013	0.011	0.034	0.023
31+	0.280	0.234	-0.046	0.006	0.013	0.007
B. Weighted Average (using alternative shares)						
Actual shares	0.136	0.183	0.047			
1973-75 shares	0.136	0.148	0.012			
2000-02 shares	0.149	0.183	0.034			

Notes: "*" indicates that the change in the variance is significantly different from zero at the 95 percent

TABLE I
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 * experience/10	−0.0032	−0.0234	−0.0068	−0.0193
	(0.0094)	(0.0123)	(0.0095)	(0.0127)
(e) Standardized AFQT * experience/10		0.0752		0.0515
		(0.0286)		(0.0343)
(f) Black * experience/10			−0.1315	−0.0834
			(0.0482)	(0.0581)
R^2	0.2861	0.2870	0.2870	0.2873

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 * experience/10	−0.0539	−0.0952	−0.0665	−0.0889
	(0.0235)	(0.0276)	(0.0234)	(0.0283)
(e) Standardized AFQT * experience/10		0.1407		0.0913
		(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

Experience is modeled with a cubic polynomial. All equations control for year effects, education interacted with a cubic time trend, Black interacted with a cubic time trend, AFQT interacted with a cubic time trend, two-digit occupation at first job, and urban residence. For these time trends, the base year is 1992. For the model in Panel 1 column (1) the coefficient on AFQT and Black are .0312 and −.1006, respectively, when evaluated for 1983. In Panel 2 the instrumental variables are the corresponding terms involving potential experience and the other variables in the model. Standard errors are White/Huber standard errors computed accounting for the fact that there are multiple observations for each worker. The sample size is 21,058 observations from 2976 individuals.

TABLE II
THE EFFECTS OF FATHER'S EDUCATION, SIBLING WAGES, AND SCHOOLING ON WAGES
Dependent Variable: Log Wage; Experience Measure: Potential Experience.
OLS estimates (standard errors)

Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(a) Education	0.0511	0.0630	0.0568	0.0659	0.0666	0.0730	0.0704	0.0734
	(0.0160)	(0.0166)	(0.0163)	(0.0167)	(0.0129)	(0.0140)	(0.0130)	(0.0140)
(b) Black	−0.2074	−0.2076	−0.0509	−0.0878	−0.2212	−0.2209	−0.0705	−0.0793
	(0.0276)	(0.0276)	(0.0846)	(0.0871)	(0.0250)	(0.0250)	(0.0668)	(0.0692)
(c) Log of sibling's wage	0.1802	−0.0260	0.1817	0.0010				
	(0.0328)	(0.0913)	(0.0329)	(0.0940)				
(d) Father's education/10					0.0826	−0.0187	0.0829	0.0314
					(0.0366)	(0.1000)	(0.0364)	(0.1030)
(e) Education * experience/10	0.0107	0.0012	0.0065	−0.0008	0.0023	−0.0029	−0.0002	−0.0027
	(0.0131)	(0.0136)	(0.0133)	(0.0136)	(0.0104)	(0.0113)	(0.0105)	(0.0113)
(f) Log of sibling's wage * experience/10		0.1796		0.1571				
		(0.0749)		(0.0770)				
(g) Father's education * experience/100						0.0867		0.0441
						(0.0813)		(0.0841)
(h) Black * experience/10			−0.1311	−0.1004			−0.1270	−0.1194
			(0.0686)	(0.0704)			(0.0541)	(0.0563)
R^2	0.3183	0.3196	0.3191	0.3200	0.2748	0.2750	0.2755	0.2756
Observations	10746	10746	10746	10746	18523	18523	18523	18523
Individuals	1441	1441	1441	1441	2594	2594	2594	2594

Experience is modeled with a cubic polynomial. All equations control for year effects, education interacted with a cubic time trend, Black interacted with a cubic time trend, two-digit occupation at first job, and urban residence. Columns (1)–(4) control for sibling's gender and the log of sibling's wage interacted with a cubic time trend. Columns (5)–(8) control for father's education interacted with a cubic time trend. For these time trends, the base year is 1992. For the models in columns (1) and (5), the coefficients on log of sibling wage and father's education are .1680 and .0357, respectively, when evaluated for 1983. Standard errors are White/Huber standard errors computed accounting for the fact that there are multiple observations for each worker.

Lange (2006)

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.