Schooling, Experience, and Earnings

J.Angrist
MIT 14.661 Fall 2024

- We learn much of what we know "on the job"
  - Assistant professors-to-be: you'll eventually learn to teach, maybe even to write (but we don't school you much in either of these skills)
- The Mincer (1974) framework puts experience and schooling together
- Becker (1964) asks: who pays for your OJT? Mincer simplifies by assuming the answer, so we start with this

Specific Ideas About General HK

- A distinction drawn using two periods, with OJT done in period 0:

$$\pi(L) = pf_0(L) + \frac{pf_1(L)}{1+r} - (w_0 + c)L - \frac{w_1}{1+r}L$$

  where c is training cost and w_0, w_1 are wages in periods 0 and 1
  - Assume $f_1'(L) > f_0'(L)$ by virtue of OJT paid for by cost c (the firm remits this, but where does the incidence lie?)
- FOC:

$$pf_0'(L) + \frac{pf_1'(L)}{1+r} = (w_0 + c) + \frac{w_1}{1+r},$$

  which we can write:

$$MP_0 + \frac{MP_1}{1+r} = (w_0 + c) + \frac{w_1}{1+r}$$

- The earnings profile
  - If training raises my value to MP_1 for all employers, my OJT is said to produce *general HK*
    * You gotta pay me what I'm worth to others: $w_1 = MP_1$, because others will pay that
  - This implies: $w_0 = MP_0 - c$, from which we conclude that *workers pay for general HK* in the form of lower initial wages

- If the HK from training is *specific*, that is, when the employer providing OJT benefits uniquely from the resulting productivity boost, the period 1 wage need only clear MP_1 elsewhere (this might be equal to pre-training productivity, MP_0)
  - Employers providing specific HK may share the costs of my OJT, paying $w_0 > MP_0 - c$ and $MP_0 < w_1 < MP_1$.
  - Paying a little more than MP_0 in period 1 ensures I stick around after training.
- As we'll soon see, the view that most HK is general (and therefore paid for by workers) can explain why wages increase with experience (though this is not the only possible explanation for that)
- A noteworthy policy implication of general HK: min wage prevents inexperienced low-skill workers from acquiring general skills (Neumark and Wascher 2001 explore this)
- Acemoglu and Pischke (1999) and Autor (2001) question the Beckerian conclusion that workers pay for general training. Many firms indeed pay for what looks like general skills training (as does the MIT econ department for its IT staff). The explanation, as with many such puzzles, revolves around worker heterogeneity and information asymmetry:
  - If training is more productive and therefore more valuable to high ability workers, firms that offer to train can lower wages initially while inducing favorable self-selection
  - Workers of high perceived ability choose firms offering training in expectation of wage gains in permanent employment, while low ability workers are deterred by lower wages and limited expected gains
  - In Autor (2001), temporary help firms garner a short-run information advantage (over their clients) that allows them to earn monopsony profits

The Mincer Earnings Function

- Mincer (1974) asks
  - How do schooling and OJT determine wages over working life?
- Assumptions
  - HK production technology is given
  - The investment path is exogenous; Ben-Porath (1967) analyzes the endogenous case, but it's not empirically very tractable; Mincer can be interpreted as choosing a specification suggested by the BP67 analysis

- All skills are general, so workers pay for skill acquisition through foregone wages

- Details
  - $k(t)$ = fraction of potential earnings devoted to investment in HK
  - Earnings are given by $y(t) = (1 - k(t))E(t)$ where $E(t)$ is potential earnings, aka my marginal product, which determines my pay if I don't spend anything on OJT
  - The rate of return on HK is fixed at ρ for all workers, a parameter determined by market forces (equalizing differences, perhaps)
  - Assuming I reap the rewards to my OJT in continuous time, we have:

$$E'(t) = \rho k(t) \cdot E(t) = g(t) \cdot E(t)$$

    This simple differential equation has solution:

$$E(t) = E(0)e^{\int_0^t g(\tau)d\tau} \tag{1}$$

- Schooling
  - For the first s years of life, set $k(\tau) = 1$, with $k(\tau) = 0$ thereafter:

$$\begin{aligned} g(\tau) = \rho k(\tau) &= \rho \qquad 0 \leq \tau \leq s \\ &= 0 \qquad otherwise \end{aligned}$$

    so for $t > s$,

$$y(t) = E(s) \equiv E(0)e^{\rho s}$$

    and

$$\ln y(t) = \ln y(0) + \rho s,$$

    as in the equalizing differences story

- Experience
  - For investment periods $\tau > s$, define work experience $x = \tau - s$ and parameterize OJT investment as:

$$\tilde{k}(x) = k_0(1 - \frac{x}{T}),$$

    for $0 \leq x \leq T$ and 0 otherwise
  - This comes from BP67 in the sense that: (a) some post-schooling investment is optimal; (b) investment should decline with age because the payoff horizon shrinks
  - Next, write earnings as a function of schooling and what comes after:

$$E(s,t) = E(0)e^{\int_0^s g(\tau)d\tau + \int_s^t g(\tau)d\tau} = E(s)e^{\int_s^t g(\tau)d\tau} \tag{2}$$

  - Change vars from t to $x \equiv t - s$:

$$E(s,x) = E(s)e^{\int_0^x \tilde{g}(\tau)d\tau} \tag{3}$$

    where $\tilde{g}(x) = \rho\tilde{k}(x)$ describes investment as a function of $x = t - s$

- Earnings
  - Finally, use $y(s,x) = [1-\tilde{k}(x)]E(s,x)$, take logs, and integrate (3) to obtain the *Mincer wage equation:*

$$\begin{aligned}\ln y(s,x) &\approx \ln E(0) + \rho s + \rho k_0 x - \frac{\rho k_0}{2T}x^2 - k_0(1-\frac{x}{T}) \\ &= [\ln E(0) - k_0] + \rho s + \left[\rho k_0 + \frac{k_0}{T}\right] x - \frac{\rho k_0}{2T}x^2\end{aligned}$$

    Key features
    - Log wages are linear in schooling and quadratic in experience
  - When do earnings peak? Solve for x^* in

$$x^* = \frac{\rho k_0 + \frac{k_0}{T}}{\rho k_0/T} = T + 1/\rho$$

    which equals 40 for $T = 30$ and $\rho = .1$
    - Yikes, I'm over the hill! (How can I avoid obsolescence?)
  - Note that schooling initially reduces earnings by reducing experience, but the cost of this fades since the experience profile is concave
- Questions
  - Why is it fair to say this model treats schooling and experience as qualitatively similar?
  - How can you tell the model assumes that OJT generates general rather than specific skills?

Earnings equation 'metrics

- Mincer functional forms put to the test in Murphy and Welch (1990)
- Experience vs seniority examined in Altonji and Shakotko (1987) and Topel (1991)
- Causality, as always, at issue

Angrist 1990: The price of service

- Draftees suffer a loss of earnings and, not coincidentally, perhaps, a loss of experience
- Enriching the draft lottery first stage (Angrist and Chen, 2011)
- Other experience experiments: plant closures and layoffs as in Jacobson, Lalonde, and Sullivan (1993)

14.64/661
Experience Profiles

[Source: Murphy & Welch (JoLE, 1990)]

FIG. 1.—Average earnings profiles from quadratic, 1963–86

Points:
1. Wages increase with education
2. Wages increase with experience
3. Wages are log-concave

log differences ~ percent differences

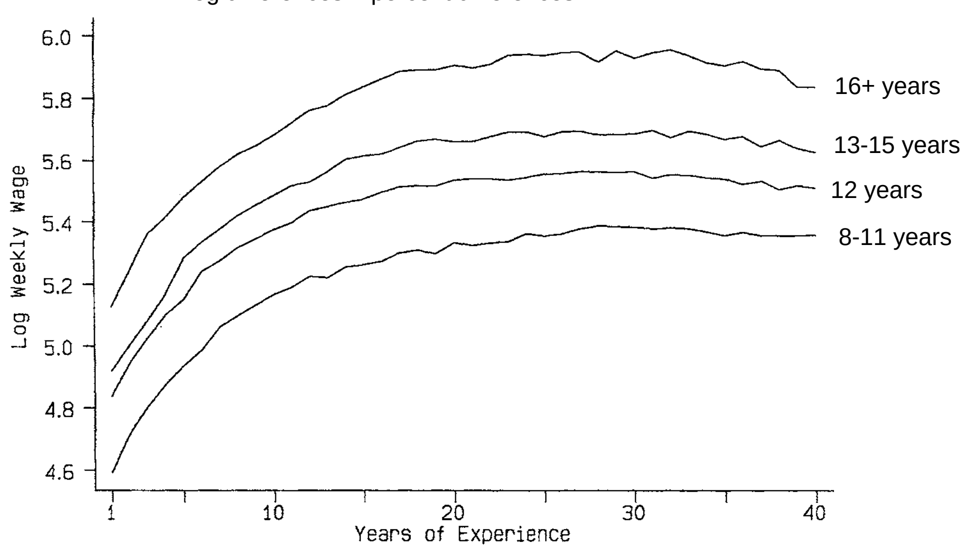


FIG. 2.—Actual average earnings profiles, 1963–86

"experience" = age - education - 6

Wald Serves in **Vietnam**

- Key variables

$$\begin{aligned} \mathrm{Z}_i &= \text{randomly assigned draft-eligibility in 1970-72 draft lotteries} \\ \mathrm{D}_i &= \text{a dummy indicating Vietnam-era veterans} \\ \mathrm{Y}_i &= \text{earnings after service} \end{aligned}$$

- The causal effect of Vietnam-era military service is the draft-eligibility RF divided by the draft-eligibility first stage
  - D_i is also a dummy, so the first stage is a diff in probs:

$$\begin{aligned} \frac{Cov(\mathrm{D}_i, \mathrm{Z}_i)}{V(\mathrm{Z}_i)} &= E[\mathrm{D}_i|\mathrm{Z}_i = 1] - E[\mathrm{D}_i|\mathrm{Z}_i = 0] \\ &= P[\mathrm{D}_i = 1|\mathrm{Z}_i = 1] - P[\mathrm{D}_i = 1|\mathrm{Z}_i = 0] \end{aligned}$$

- The **RF** is a diff in means
- **Angrist (1990), Figures 1-2 and** Table 3

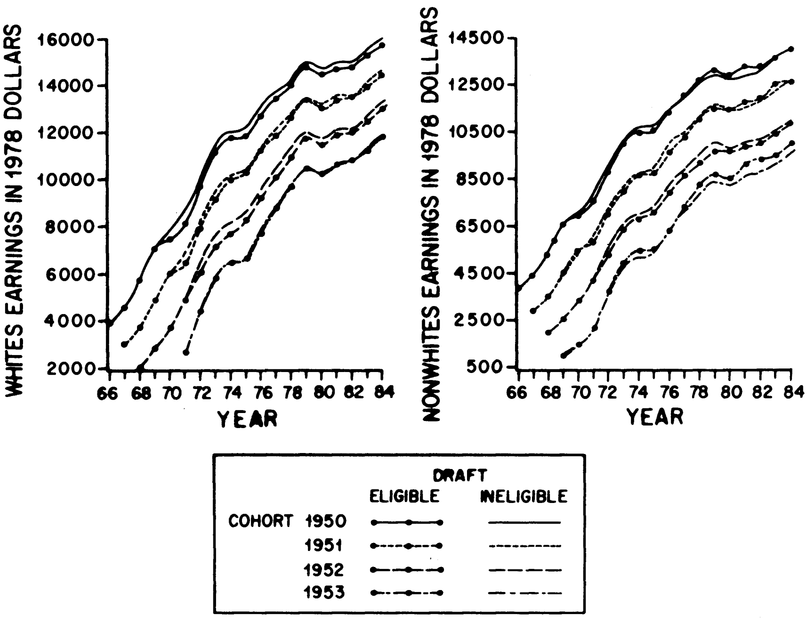


FIGURE 1. SOCIAL SECURITY EARNINGS PROFILES BY DRAFT-ELIGIBILITY STATUS

Notes: The figure plots the history of FICA taxable earnings for the four cohorts born 1950–3. For each cohort, separate lines are drawn for draft-eligible and draft-ineligible men. Plotted points show average real (1978) earnings of working men born in 1953, real earnings +\$3000 for men born in 1950, real earnings +\$2000 for men born in 1951, and real earnings +\$1000 for men born in 1952.

TABLE 3—WALD ESTIMATES

Cohort	Year	Draft-Eligibility Effects in Current $: FICA Earnings (1)	Adjusted FICA Earnings (2)	Total W-2 Earnings (3)	$\hat{p}^e - \hat{p}^n$ (4)	Service Effect in 1978 $ (5)
1950	1981	−435.8	−487.8	−589.6	0.159	−2,195.8
		(210.5)	(237.6)	(299.4)	(0.040)	(1,069.5)
	1982	−320.2	−396.1	−305.5		−1,678.3
		(235.8)	(281.7)	(345.4)		(1,193.6)
	1983	−349.5	−450.1	−512.9		−1,795.6
		(261.6)	(302.0)	(441.2)		(1,204.8)
	1984	−484.3	−638.7	−1,143.3		−2,517.7
		(286.8)	(336.5)	(492.2)		(1,326.5)
1951	1981	−358.3	−428.7	−71.6	0.136	−2,261.3
		(203.6)	(224.5)	(423.4)	(0.043)	(1,184.2)
	1982	−117.3	−278.5	−72.7		−1,386.6
		(229.1)	(264.1)	(372.1)		(1,312.1)
	1983	−314.0	−452.2	−896.5		-2,181.8
		(253.2)	(289.2)	(426.3)		(1,395.3)
	1984	−398.4	−573.3	−809.1		−2,647.9
		(279.2)	(331.1)	(380.9)		(1,529.2)
1952	1981	−342.8	−392.6	−440.5	0.105	−2,502.3
		(206.8)	(228.6)	(265.0)	(0.050)	(1,556.7)
	1982	−235.1	−255.2	−514.7		−1,626.5
		(232.3)	(264.5)	(296.5)		(1,685.8)
	1983	−437.7	−500.0	−915.7		−3,103.5
		(257.5)	(294.7)	(395.2)		(1,829.2)
	1984	−436.0	−560.0	−767.2		−3,323.8
		(281.9)	(330.1)	(376.0)		(1,959.3)

Notes: Standard errors in parentheses.
Columns (1) and (3) are taken from Table 1.
Column (2) reports draft-eligibility treatment effects on earnings adjusted for censoring at the FICA taxable maximum. The adjustment procedure is described in the Appendix. Column (4) reports SIPP estimates of the effect of draft eligibility on veteran status, taken from Table 2. Column (5) reports estimates of the effect of military service on civilian earnings is implied by columns (2) and (4).

Multiple groups and 2SLS

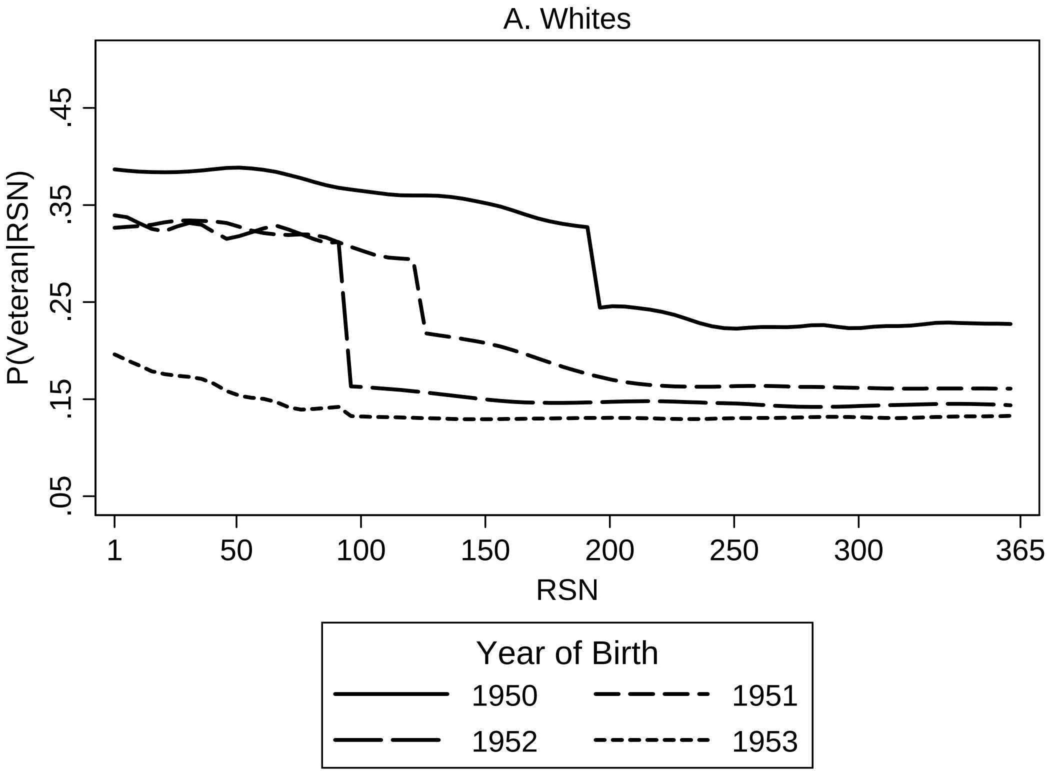
- More to the draft lottery than draft-eligibility: **Angrist and Chen (2011), Figure 1**
- Let $\mathrm{R}_i = j \in \{1, ..., J\}$ denote lottery numbers. Draft-eligibility Wald uses $1[\mathrm{R}_i < 195]$ as an instrument in a just-identified setup
- Using fine-grained info on R_i, we have

$$E[\mathrm{Y}_i|\mathrm{R}_i] = \alpha + \rho P[\mathrm{D}_i = 1|\mathrm{R}_i], \tag{11}$$

since $P[\mathrm{D}_i = 1|\mathrm{R}_i] = E[\mathrm{D}_i|\mathrm{R}_i]$. So we can estimate ρ by fitting:

$$\bar{y}_j = \alpha + \rho\hat{p}_j + \bar{\eta}_j;\ j = 1, ..., J \tag{12}$$

- Efficient GLS for this grouped constant-effects linear model is weighted least squares, weighted by $V(\bar{\eta}_j)$
  - $V(\bar{\eta}_j) = \frac{\sigma_\eta^2}{n_j}$ under homoskedasticity



A. Whites
P(Veteran|RSN)
.05
.15
.25
.35
.45
1
50
100
150
200
250
300
365
RSN
Year of Birth
1950
1951
1952
1953

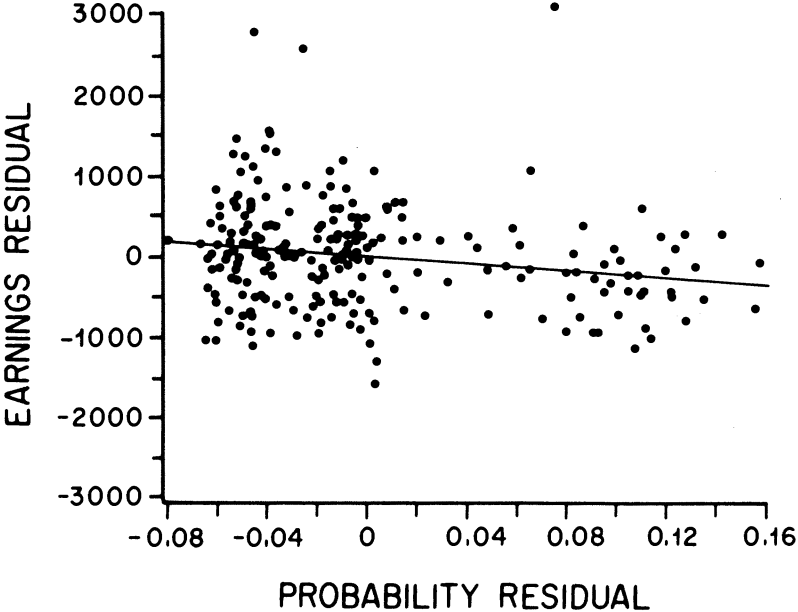


FIGURE 3. EARNINGS AND THE PROBABILITY OF VETERAN STATUS BY LOTTERY NUMBER

Notes: The figure plots mean W-2 compensation in 1981–4 against probabilities of veteran status by cohort and groups of five consecutive lottery numbers for white men born 1950–3. Plotted points consist of the average residuals (over four years of earnings) from regressions on period and cohort effects. The slope of the least-squares regression line drawn through the points is $-2{,}384$, with a standard error of 778, and is an estimate of α in the equation

$$\bar{y}_{ctj} = \beta_c + \delta_t + \hat{p}_{cj}\alpha + \bar{u}_{ctj}.$$

human capital. The earnings function motivated by the theory of human capital is loglinear in years of schooling and log-quadratic in years of labor market experience. This functional form puts testable restrictions on the time-series of earnings differences by veteran status.[13]

Adapting the human capital earnings function for the problem at hand, the earnings of individual i in cohort c at time t may be written

$$\text{(4a)} \qquad y_{cti} = \delta_t + w_i\delta_0 + \beta_0 x_{ict} + \gamma x_{ict}^2 + u_{it},$$

where y_{cti} now denotes log earnings, δ_t is a time-varying intercept, β_0, γ, and δ_0 are parameters. x_{ict} is the civilian labor market experience of man i in cohort c at time t, taken here to be equal to $[t-(c+18)-w_i-s_i l]$, where w_i is the deviation of i's schooling from the sample mean level of schooling, and l is years of military experience for veterans. As before, s_i is a dummy variable that indicates military service.

To focus on parameters that can be estimated using Social Security data, equation (4) is rewritten as

$$\text{(4b)} \qquad y_{cti} = \delta_{it} + \beta_0(x_{ct} - s_i l) + \gamma(x_{ct} - s_i l)^2 - (2\gamma w_i x_{ct} - 2\gamma l w_i s_i) + u_{it},$$

where $x_{ct} = t-(c+18)$ and $\delta_{it} = \delta_t + w_i(\delta_0 - \beta_0) + \gamma w_i^2$. Now, as in the previous analysis, assume that schooling does not vary by lottery number. Assume also that schooling is independent of cohort—this seems reasonable for the small cohort range considered here. Finally, to focus solely on the loss of labor market experience, assume that schooling is independent of veteran status. Then using dummy instrumental variables to group equation (4b) by cohort, year, and lottery number, average log earnings for members of cohort c at time t in lottery-number cell j are

$$\text{(5)} \qquad \bar{y}_{ctj} = \delta_t + \beta_0 x_{ct} + \gamma x_{ct}^2 - [\beta_0 l - \gamma l^2]\hat{p}_{cj} - [2\gamma l](\hat{p}_{cj} x_{ct}) + \bar{u}_{ct},$$

where δ_t now includes the period mean of δ_{it}.[14]

A generalization of model (5) allows the linear experience term to vary with veteran status by letting the slope for individual i be $\beta_i = \beta_0 + \beta_1 s_i$. In this case, mean cell earnings are characterized by

$$\text{(6)} \qquad \bar{y}_{ctj} = \delta_t + \beta_0 x_{ct} + \gamma x_{ct}^2 - [\beta_0 l - \gamma l^2 + \beta_1 l]\hat{p}_{cj} - [2\gamma l - \beta_1](\hat{p}_{cj} x_{ct}) + \bar{u}_{ct}.$$

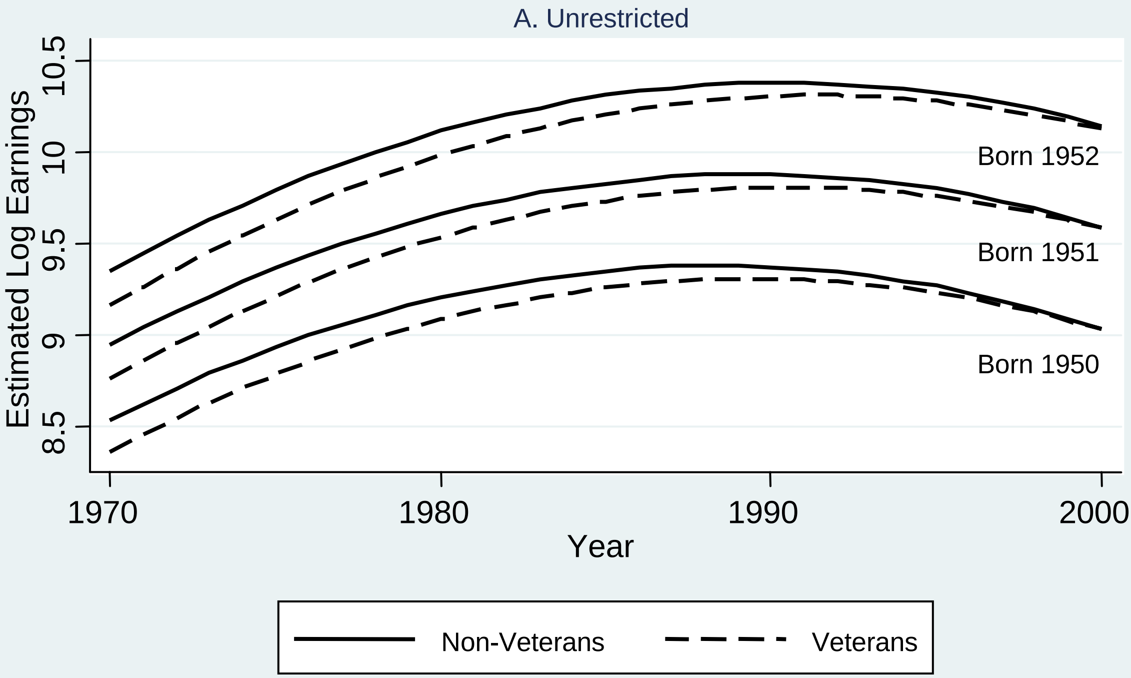
Models (5) and (6) both have the following reduced form in terms of unrestricted regression coefficients:

$$\text{(7)} \qquad \bar{y}_{ctj} = \delta_t + \beta_0 x_{ct} + \gamma x_{ct}^2 + \pi_1 \hat{p}_{cj} + \pi_2(\hat{p}_{cj}^* x_{ct}) + \bar{u}_{ct}.$$

Note that the reduced form veteran effect is $\alpha_{ct} \equiv \pi_1 + \pi_2 x_{ct}$. Thus, these models parameterize a time-varying veteran status coefficient as a linear function of labor market

[13]See Mincer (1974) for theoretical justification of the human capital earnings function. A recent survey of the human capital literature is Willis (1986).

[14]Averaging over c, t, and j eliminates $(\partial\gamma w_i x_{ct} - \partial\gamma w_i s_i)$ because w_i is orthogonal to x_{ct} and s_i by assumption. Using the fact that $E(s_i|c,j) = E(s_i^2|c,j) = p_{cj}$, (4) simplifies to (5). Note that (5), (6), and (7) are not estimable if allowance need be made for cohort as well as period effects. Qualitatively similar estimates to those reported below were obtained when δ_t was dropped in favor of cohort effects, although the goodness-of-fit test leads to rejection of models without period effects.



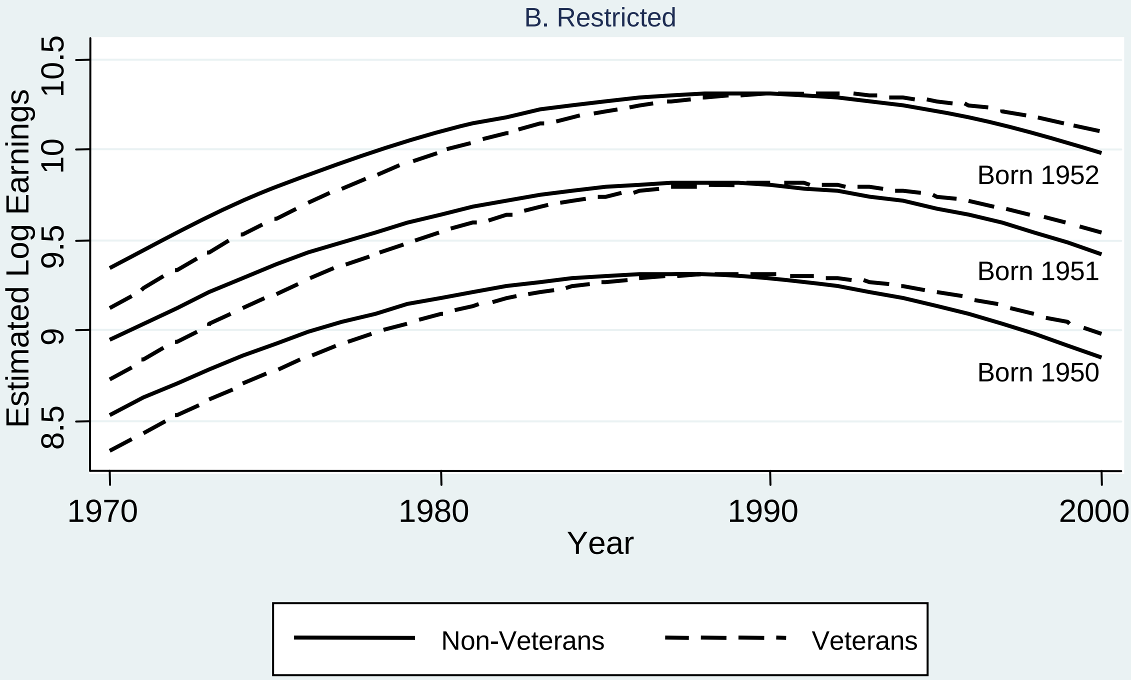


Figure 2. The Effect of Veteran Status on Experience Profiles

TABLE 5—EARNINGS-FUNCTION MODELS FOR THE VETERAN EFFECT, WHITES BORN 1950–52

Parameter	Model (5): Loss of Experience (1)	Model (6): Loss of Experience, Reduced Growth Rate (2)	Model (7): Unrestricted Reduced Form (3)
Experience Slope, β_0	0.1022	0.1016	0.1016
	(0.007)	(0.007)	(0.007)
Experience Squared, γ	−0.0027	−0.0025	−0.0025
	(0.0003)	(0.0003)	(0.0003)
Veteran Effect on Slope, β_1		−0.0035	
		(0.0023)	
Veteran Loss of Experience, l	2.08	1.84	
	(0.38)	(0.43)	
$\pi_1 = -[\beta_0 l - \gamma l^2 + \beta_1 l]$			−0.189
			(0.052)
$\pi_2 = -[2\gamma l - \beta_1]$			0.006
			(0.004)
Age at Which Reduced Form Veteran Effect $(\pi_1 + \pi_2 x_{ct}) = 0$			50.1
			(15.9)
χ^2(dof)	1.41(1)		813.57(1247)

Notes: Standard errors in parentheses.
The table reports estimates of experience-earnings profiles that include parameters for the effect of veteran status. Estimates are of equations (5), (6) and (7) in the text. The estimating sample includes FICA taxable earnings from 1975–84 for men born 1950, 1976–84 earnings for men born in 1951, and 1977–84 earnings for men born 1952. The estimation method is optimally weighted Two-Sample Instrumental Variables for a nonlinear model in columns (1) and (2), and for a linear model in column (3).