EXHIBIT C

Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records

By Joshua D. Angrist*

The randomly assigned risk of induction generated by the draft lottery is used to construct estimates of the effect of veteran status on civilian earnings. These estimates are not biased by the fact that certain types of men are more likely than others to service in the military. Social Security administrative records indicate that in the early 1980s, long after their service in Vietnam was ended, the earnings of white veterans were approximately 15 percent less than the earnings of comparable nonveterans. (JEL 824)

A central question in the debate over military manpower policy is whether veterans are adequately compensated for their service. The political process clearly reflects the desire to compensate veterans: since World War II, millions of veterans have enjoyed benefits for medical care, education and training, housing, insurance, and job placement. Recent legislation provides additional benefits for veterans of the Vietnam era. Yet, academic research has not shown conclusively that Vietnam (or other) veterans are worse off economically than nonveterans. Many studies find that Vietnam veterans earn less than nonveterans, but others find positive effects, or effects that vary with age and schooling. Regarding the general position of veterans, a member of the Twentieth Century Fund's Task Force on Policies Toward Veterans concludes that "Within any age group, veterans have higher incomes, more education, and lower unemployment rates than their nonveteran counterparts."[1]

The goal of this paper is to measure the long-term labor market consequences of military service during the Vietnam era. Previous research comparing civilian earnings by veteran status may be biased by the fact that certain types of men are more likely to serve in the armed forces than others. For example, men with relatively few civilian opportunities are probably more likely to enlist. Estimation strategies that do not control for differences in civilian earnings potential will incorrectly attribute lower civilian earnings of veterans to military service. The research reported here overcomes such statistical problems by using the Vietnam era draft

*Department of Economics, Harvard University, Cambridge, MA 02138. Grateful thanks go to Warren Buckler, Cresston Smith, Ada Enis, and Bea Matsui for their assistance in producing the Social Security data; to Chester Bowie for his help in producing the SIPP data; and to Mike Dove for providing DMDC administrative records. Special thanks also go to David Card and Whitney Newey, from whose instruction and comments I have benefited greatly, and to Alan Krueger and an anonymous referee, whose careful reviews of an earlier draft led to substantial improvement. Data collection for this project was funded by the Princeton Industrial Relations Section. Funds for computation and financial support of the author were provided by the Industrial Relations Section, the Princeton Department of Economics, the Sloan Foundation, and the Olin Foundation.

[1]The quote is from Michael Taussig (1974, p. 51). Legislation pertaining to veterans benefits is outlined in Veterans Administration (1984) and in other annual reports of the Veterans Administration. Studies by Sherwin Rosen and Paul Taubman (1982), Saul Schwartz (1986), and Jon Crane and David Wise (1987) find that Vietnam veterans earn less than nonveterans. Dennis DeTray (1982) and Mark Berger and Barry Hirsch (1983) find some positive effects for different age and schooling classes, and Veterans Administration (1981a) researchers find an overall positive effect.

lotteries to set up a natural experiment that randomly influenced who served in the military.[2]

Section I describes the Social Security administrative records used in the empirical work and provides background on the draft lotteries. In each lottery, priority for induction was determined by a Random Sequence Number (RSN) from 1–365 that was assigned to birthdates in the cohort being drafted. Men were called for induction by RSN up to a ceiling determined by the Defense Department, and only men with lottery numbers below the ceiling could have been drafted. Therefore, men with lottery numbers below the ceiling are referred to here as "draft-eligible."

The empirical analysis begins in Section II with estimates of the effect of draft eligibility on earnings. If draft eligibility is correlated with veteran status but uncorrelated with other variables related to earnings, then earnings differences by draft-eligibility status can be attributed to military service. In Section III, information on the proportions of draft-eligible and draft-ineligible men who actually served in the military is used to convert estimates of the effect of draft eligibility into estimates of the effect of military service. The assumptions underlying this procedure are those that justify instrumental variables estimation; in principle, any function of the RSN provides a legitimate instrument for veteran status. In the second part of Section III, an instrumental variables estimation strategy is developed which is more efficient than one based solely on draft-eligibility status. Results in Section III indicate that white veterans earn approximately 15 percent less than nonveterans as much as ten years after their discharge from the military.

Section IV tests the hypothesis that veterans earn less than nonveterans because they have less civilian labor market experience. Results in this section suggest that the earnings loss to white veterans is equivalent to a loss of two years of civilian labor market experience. Section V reviews some of the potential pitfalls in estimation based on the draft lottery. Section VI offers conclusions and indicates directions for future research.

I. Background and Data

A. *National Random Selection*[3]

There were five draft lotteries during the Vietnam War period. The 1970 lottery covered 19- to 26-year-old men born in 1944–50, although most of the men drafted in 1970 were born in 1950. Other lotteries were restricted to 19- and 20-year-olds. The 1971 lottery covered men born in 1951, the 1972 lottery covered men born in 1952, and so on, through 1975. However, no one was drafted after 1972, and congressional conscription authority expired in July 1973.

Draft lottery RSNs were randomly assigned in a televised drawing held a few months before men reaching draft age were to be called.[4] Draft-eligibility ceilings—RSN 195 in 1970, RSN 125 in 1971, and RSN 95 in 1972—were announced later in the year, once Defense Department manpower needs were known. As a consequence of this delay, many men with low numbers volunteered for the military to avoid being drafted and to improve their terms of service (Angrist 1989b). There was even a behavioral response to the lottery in enlistment rates for the 1953 cohort, although no one born in 1953 was drafted. In the analysis that follows, the "draft-eligibility ceiling" for men born in 1953 is set at RSN 95, the highest lottery number called in 1972.

[2]A candid assessment of the problems caused by nonrandom selection for military service is given by Crane and Wise (1987), who note they were unable to use econometric sample selection models to generate robust estimates of the effects of military service on civilian earnings. The first researchers to use the lottery to solve the selection problem were Norman Hearst, Tom Newman, and Stephen Hulley (1986), who present lottery-based estimates of delayed effects of military service on mortality.

[3]This section draws on Curtis Tarr (1981) and the Selective Service System (1986).

[4]Men born from 1944–49 were already of draft age when the 1970 lottery was held on December 1, 1969. For nonveterans in this group, subsequent liability for service was determined by 1970 lottery numbers.

Only the initial selection process was based on RSN order. Subsequent selection from the draft-eligible, nondeferred pool was based on a number of criteria. The most important screening criteria were the pre-induction physical examination and a mental aptitude test. In 1970, for example, half of all registrants failed pre-induction examinations and 20 percent of those who passed were eliminated by physical inspections conducted at induction (Selective Service System, 1971). Of course, the fact that armed forces selection criteria were ultimately not random does not mean that the initial *priority* for induction was not randomly assigned by RSN.

The year 1970 was the last time men over the age of 20 were drafted. In principle, nonveterans born between 1944 and 1949 continued to be at risk of induction in the 1970 lottery, but the majority of men who ended up serving from these cohorts had already entered the military by the time of the 1970 lottery drawing. Veterans born from 1944–49 who managed to avoid service until 1970 may not constitute a representative sample. Therefore, the analysis here is restricted to men who turned 19 in the year they were at risk of induction. This sample includes men who were born between 1950 and 1953.

B. *Social Security Earnings Data*

Earnings data used in this study are drawn from the Social Security Administration's (SSA) Continuous Work History Sample (CWHS). The CWHS data set, described in detail in the Appendix, is a one percent sample drawn from all possible Social Security numbers. The CWHS includes two earnings series: the first contains information on the 1964–84 earnings of men in employment covered by FICA (Social Security) up to the Social Security taxable maximum. It also includes FICA taxable earnings from self-employment. The second series, beginning in 1978, contains total compensation as reported on Internal Revenue Service Form W-2, excluding earnings from self-employment. In principle, the W-2 earnings data are neither censored nor limited to earnings from Social Security taxable employment. However, because SSA procedures for the collection of W-2 forms are relatively new, W-2 earnings data are probably less reliable than the FICA data.

The original CWHS data set does not contain information on date of birth. SSA programmers matched date of birth variables to the CWHS in a special extract created for this project. Lottery numbers were then matched to dates of birth, using tables published in the 1969–73 Semiannual Reports of the Director of Selective Service.

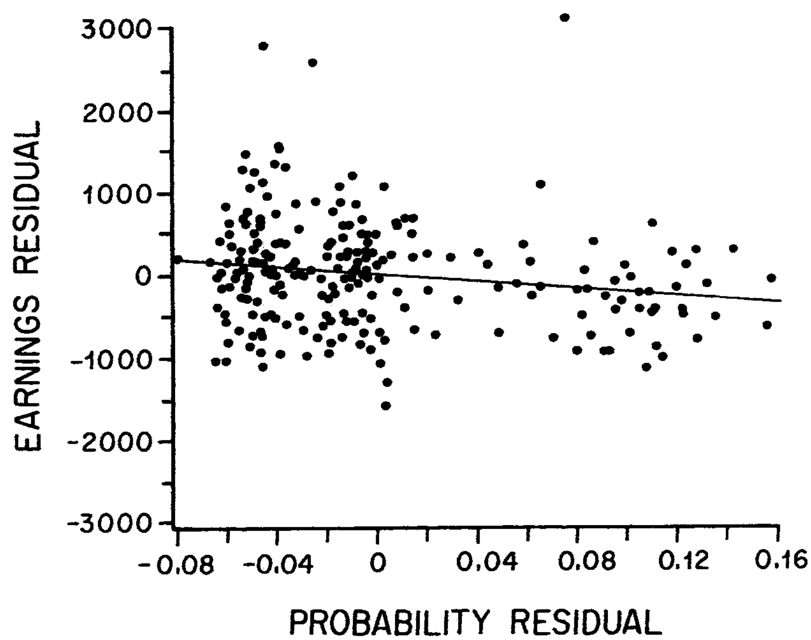
The Internal Revenue Service limits disclosure of data collected for tax purposes. To adhere to these disclosure requirements, the SSA could release only aggregate data. The aggregate data set contains sample statistics for cells defined by year of earnings, year of birth, race, and five consecutive lottery numbers. Cell statistics include means, variances, fraction with earnings equal to the taxable maximum, fraction with earnings above the taxable maximum, fraction with zero earnings, and number of observations in each cell.

II. The Effect of Draft Eligibility on Earnings

Figure 1 shows the history of FICA taxable earnings for draft lottery participants born between 1950 and 1953.[5] For each cohort there are two lines drawn: one for draft-eligible men, and one for men with lottery numbers that exempted them from the draft.

The impact of draft eligibility on the earnings profiles is striking. There appears to be no difference in earnings until the year of conscription risk in the draft lottery. Subsequently, the earnings of draft-eligible white men born in 1950–52 fall below the earnings of draft-ineligible white men born in 1950–52. The earnings of draft-eligible nonwhites also fall below the earnings of other non-

[5]Earnings are in 1978 dollars. The deflator used for all tabulations is the CPI on p. 313 of *The Economic Report of the President* (Council of Economic Advisors, 1988).



Notes: The figure plots the history of FICA taxable earnings for the four cohorts born 1950–53. 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.

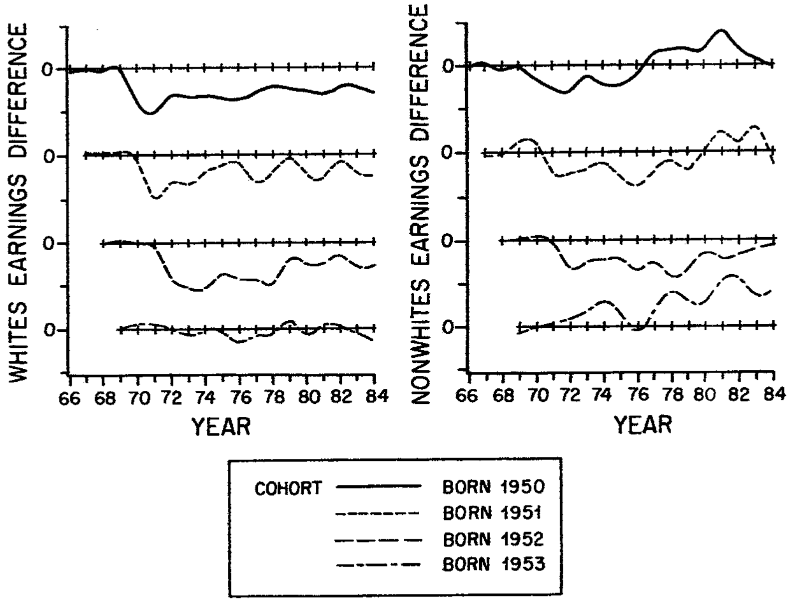
FIGURE 1. SOCIAL SECURITY EARNINGS PROFILES BY DRAFT-ELIGIBILITY STATUS

whites, but the gap appears to narrow and become positive for some nonwhite cohorts in later years. The fact that earnings do not differ by draft-eligibility status before the lotteries is a consequence of the random assignment of draft eligibility. The only thing that distinguishes draft-eligible men from draft-ineligible men is the higher conscription risk faced by eligible men after the lottery.

Figure 2 presents a magnified view of the effect of draft eligibility on earnings. This figure plots the time-series of differences in earnings by draft-eligibility status for each cohort. As in Figure 1, Figure 2 shows no difference during the years before the year of conscription risk, while in subsequent years, the earnings histories diverge. Figure 2 also shows that the loss of earnings to draft-eligible white men was largest during the period they were most likely to be in the service. However, the earnings of draft-eligible white men continued to lag behind the earnings of draft-ineligible white men through 1984.

The picture for nonwhites is less clear. The earnings of draft-eligible nonwhites born in 1950 and 1951 exceed those of draft-ineligible nonwhites in some of the later years. On the other hand, for nonwhites born in 1952, time-series variation in earnings differences by eligibility status is similar to that of whites. The general impression for the three older cohorts of nonwhites is that the earnings of draft-eligible men at least had caught up with the earnings of draft-ineligible men by 1984.

Earnings of white men born in 1953 do not appear to differ by draft-eligibility status. The earnings of draft-eligible nonwhites born in 1953 generally exceed the earnings of nonwhites who were not draft-eligible. Differences between the effect of draft eligi-



Notes: The figure plots the difference in FICA taxable earnings by draft-eligibility status for the four cohorts born 1950–53. Each tick on the vertical axis represents $500 real (1978) dollars.

FIGURE 2. THE DIFFERENCE IN EARNINGS BY DRAFT-ELIGIBILITY STATUS

bility on men born in 1953 and the effect on the three older cohorts might be explained by the transition to an All-Volunteer Force in 1973. Men who volunteer for the military are probably less likely than draftees to suffer a career disadvantage from their service.

Estimates of the effect of draft eligibility are reported in Table 1 for both FICA earnings and W-2 earnings. Standard errors associated with the estimates are reported in parentheses. The statistics in Table 1 show that the loss in FICA earnings to draft-eligible white men is sometimes statistically significant and amounts to 2–3 percent of earnings. Estimated W-2 earnings losses are similar, but tend to be larger and more variable than the estimated losses in FICA earnings. In contrast, differences in earnings by draft-eligibility status for nonwhites rarely exceed their standard errors.

Elsewhere (Angrist 1989c), I have shown that draft-eligible white men are less likely to have earnings above the FICA taxable maximum than draft-ineligible white men. The effect of draft eligibility on nonwhites' probability of being at the taxable maximum, although imprecisely measured, also appears to go in the same direction as the effect of draft eligibility on mean earnings. These results are worth noting because, when the effect of draft eligibility on the probability of being censored has the same sign as the effect on earnings, estimates tabulated using censored data tend to underestimate the true effect.[6]

[6]The effect of censoring on estimated treatment effects is discussed in the appendix. Angrist (1989c) also reports estimates of the effect of draft eligibility on the probability of having no recorded earnings. These tabulations indicate that draft-eligible whites were somewhat more likely to have had FICA earnings during the years in which they were in the service, and that draft-eligible nonwhites are more likely to have had no earnings in recent years. There is no statistically significant evidence for either race, however, of any lasting effect of draft eligibility on the probability of having zero earnings.

TABLE 1—DRAFT-ELIGIBILITY TREATMENT EFFECTS FOR EARNINGS

	Whites							
	FICA Taxable Earnings				Total W-2 Compensation			
Year	1950	1951	1952	1953	1950	1951	1052	1953
66	−21.8							
	(14.9)							
67	−8.0	13.1						
	(18.2)	(16.4)						
68	−14.9	12.3	−8.9					
	(24.2)	(19.5)	(19.2)					
69	−2.0	18.7	11.4	−4.0				
	(34.5)	(26.4)	(22.7)	(18.3)				
70	−233.8	−44.8	−5.0	32.9				
	(39.7)	(36.7)	(29.3)	(24.2)				
71	−325.9	−298.2	−29.4	27.6				
	(46.6)	(41.7)	(40.2)	(30.3)				
72	−203.5	−197.4	−261.6	2.1				
	(55.4)	(51.1)	(46.8)	(42.9)				
73	−226.6	−228.8	−357.7	−56.5				
	(67.8)	(61.6)	(56.2)	(54.8)				
74	−243.0	−155.4	−402.7	−15.0				
	(81.4)	(75.3)	(68.3)	(68.1)				
75	−295.2	−99.2	−304.5	−28.3				
	(94.4)	(89.7)	(85.0)	(79.6)				
76	−314.2	−86.8	−370.7	−145.5				
	(106.6)	(102.9)	(98.2)	(93.0)				
77	−262.6	−274.2	−396.9	−85.5				
	(117.9)	(112.2)	(111.1)	(107.1)				
78	−205.3	−203.8	−467.1	−65.3	1,059.3	233.2	175.3	−1,974.5
	(132.7)	(127.0)	(127.3)	(123.1)	(2,159.3)	(1,609.4)	(1,567.9)	(912.1)
79	−263.6	−60.5	−236.8	89.2	−1,588.7	523.6	−580.8	−557.9
	(160.5)	(152.3)	(153.9)	(148.7)	(1,575.6)	(1,590.5)	(736.7)	(750.1)
80	−339.1	−267.9	−312.1	−93.8	−1,028.1	85.6	−581.3	−428.7
	(183.2)	(175.3)	(178.2)	(170.7)	(756.8)	(599.8)	(309.1)	(341.5)
81	−435.8	−358.3	−342.8	34.3	−589.6	−71.6	−440.5	−109.5
	(210.5)	(203.6)	(206.8)	(199.0)	(299.4)	(423.4)	(265.0)	(245.2)
82	−320.2	−117.3	−235.1	29.4	−305.5	−72.7	−514.7	18.7
	(235.8)	(229.1)	(232.3)	(222.6)	(345.4)	(372.1)	(296.5)	(281.9)
83	−349.5	−314.0	−437.7	−96.3	−512.9	−896.5	−915.7	30.1
	(261.6)	(253.2)	(257.5)	(248.7)	(441.2)	(426.3)	(395.2)	(318.1)
84	−484.3	−398.4	−436.0	−228.6	−1,143.3	−809.1	−767.2	−164.2
	(286.8)	(279.2)	(281.9)	(272.2)	(492.2)	(380.9)	(376.0)	(366.0)

III. The Effect of Military Service on Earnings

A. *Estimates Using Draft Eligibility*

Estimates of the effects of military service are based on a simple linear model for earnings. Denote the earnings of man i in cohort c at time t by y_{cti}, and let s_i be an indicator of veteran status. Then we may write

$$(1) \qquad y_{cti} = \beta_c + \delta_t + s_i\alpha + u_{it},$$

where β_c is a cohort effect, δ_t is a period effect common to all cohorts, and u_{it} is a residual. The coefficient α is the effect of military service on civilian earnings. If s_i is correlated with the unobserved components of the earnings equation, then α will not be consistently estimated by Ordinary Least Squares (OLS). For example, correlation between s_i and u_{it} may arise because the armed forces' eligibility criteria are correlated with earnings, but not accounted for by the econometrician, or because veterans are

TABLE 1—CONTINUED

Nonwhites

Year	FICA Taxable Earnings 1950	1951	1952	1953	Total W-2 Compensation 1950	1951	1952	1953
66	−11.8							
	(27.6)							
67	12.9	−4.0						
	(34.2)	(30.6)						
68	−29.5	−6.2	−12.0					
	(44.5)	(37.3)	(35.0)					
69	−5.1	67.8	3.4	−42.4				
	(66.8)	(53.4)	(43.4)	(36.4)				
70	−99.8	62.2	24.7	−9.0				
	(78.5)	(75.7)	(62.2)	(44.9)				
71	−164.8	−144.3	−25.0	18.2				
	(92.7)	(86.4)	(85.1)	(60.7)				
72	−188.8	−156.7	−208.2	60.4				
	(113.6)	(105.7)	(104.2)	(92.8)				
73	−85.7	−134.8	−175.6	115.5				
	(137.7)	(127.0)	(129.0)	(119.4)				
74	−179.3	−96.7	−181.4	216.5				
	(165.0)	(160.1)	(155.6)	(145.1)				
75	−190.3	−236.1	−183.7	111.6				
	(189.3)	(186.8)	(185.8)	(166.9)				
76	−105.3	−333.7	−308.9	−46.4				
	(214.7)	(215.4)	(216.5)	(199.3)				
77	112.4	−206.8	−251.1	153.5				
	(238.5)	(240.4)	(248.5)	(233.5)				
78	163.6	−108.6	−424.9	381.9	−1,145.0	2,978.2	−4,676.2	−482.7
	(272.6)	(269.2)	(279.4)	(275.7)	(2,395.6)	(2,869.6)	(1,393.1)	(2,206.0)
79	187.0	−210.3	−391.7	312.0	4,005.4	1,545.0	−494.7	−1,043.3
	(317.2)	(323.0)	(324.8)	(326.3)	(2,721.2)	(2,191.1)	(2,683.8)	(1,660.2)
80	203.2	4.8	−212.6	344.0	790.2	376.4	−292.7	288.6
	(363.1)	(368.4)	(372.5)	(370.3)	(648.1)	(533.6)	(440.9)	(416.4)
81	534.5	313.2	−305.8	717.8	802.5	415.9	−272.3	784.4
	(413.5)	(419.1)	(429.1)	(433.7)	(524.6)	(745.1)	(492.8)	(503.1)
82	285.1	175.4	−262.5	810.4	326.0	−244.3	−160.2	675.1
	(461.2)	(471.6)	(476.7)	(486.3)	(608.9)	(647.8)	(590.0)	(564.1)
83	96.0	419.5	−177.3	543.6	315.4	254.3	−53.6	462.3
	(512.6)	(538.1)	(531.5)	(523.2)	(720.0)	(767.5)	(643.4)	(638.9)
84	−76.8	−223.1	−123.3	641.3	−287.4	−718.6	−288.0	827.3
	(548.2)	(562.8)	(568.5)	(568.2)	(804.0)	(771.5)	(721.0)	(716.8)

Notes: Standard errors in parentheses.
The table shows the difference in earnings by lottery-determined draft-eligibility status.
Eligibility ceilings are RSN 195 for men born in 1950, RSN 125 for men born in 1951, and RSN 95 for men born in 1952 and 1953.
Earnings data are from the Social Security Administration CWHS, described in the text and the Appendix.

self-selected on the basis of unobserved characteristics.

The draft lottery facilitates estimation of (1) because functions of randomly assigned lottery numbers provide instrumental variables that are correlated with s_i, but orthogonal to the error term, u_{it}. For example, one such instrument is a dummy variable, d_i, that equals one if the ith individual was draft eligible. Suppose that attention is restricted to a single cohort. Then, use of d_i and a constant as instrumental variables leads to the following estimator for α:

$$(2) \qquad \hat{\alpha} = (\bar{y}^e - \bar{y}^n)/(\hat{p}^e - \hat{p}^n),$$

where $\hat{p}$ is the proportion of the cohort actually entering the military, $\bar{y}$ is mean earnings, and superscript e and superscript n denote the draft-eligible and draft-ineligible samples. Note that the numerator of (2) consists of estimates of the effect of draft eligibility plotted in Figure 2.

Intuitively, equation (2) simply adjusts earnings differences by draft-eligibility status for the fact that not all draft-eligible men actually served in the military, while some men who were not draft eligible voluntarily enlisted for service. The justification for estimation of the effects of military service in this manner is clear: it is assumed that nothing other than differences in the probability of being a veteran is responsible for differences in earnings by draft-eligibility status. This formula may also be recognized as an application of Abraham Wald's (1940) grouping method, where the data have been grouped by draft-eligibility status. Applications of this formula will therefore be referred to here as "Wald estimates."

In addition to draft-eligibility treatment effects, implementation of the Wald estimator requires estimates of $\hat{p}^e$ and $\hat{p}^n$. These estimates are tabulated from a special version of the 1984 Survey of Income and Program Participation (SIPP). The SIPP data used here were matched to an indicator of draft-eligibility status from information on birthdates included in the Census Bureau's in-house version of the SIPP file. Additional details on the SIPP data are provided in Section 7 of the Appendix.

In the upper panel of Table 2, the columns labeled $\hat{p}^e$ and $\hat{p}^n$ show probabilities of veteran status tabulated using the SIPP. Because of the small number of observations available for single-cohort statistics, each SIPP estimate is actually the average for three consecutive cohorts. For example, SIPP estimates assigned to men born in 1951 are based on data for men born in 1950, 1951, and 1952. The last column of Table 2, labeled $\hat{p}^e - \hat{p}^n$, shows the difference in the probability of military service by draft-eligibility status. Estimates of the effect of draft eligibility on veteran status for whites born 1950–52 range from 0.10 to 0.16. Thus, a rule of thumb for conversion of draft-eligibility treatment effects into estimates of the effects of military service is to multiply by $1/0.15 = 6\ 2/3$.

Wald estimates of the effect of military service for selected cohorts and years are presented in Table 3. The sample is restricted to the subset of whites born 1950–52 because the results in Table 2 suggest that this is the group for whom draft eligibility is most likely to be a useful instrument. Earnings variables are for 1981–84 because the impact of military service in these years represents a long-term effect. Furthermore, as a practical matter, both FICA and W-2 earnings data are likely to be more reliable in recent years—the FICA data because of increased employment coverage and the W-2 data because of improvements in data collection procedures.

Table 3 reports three sets of estimated draft-eligibility effects for use in the numerator of the Wald estimator. Column (1) presents estimates for FICA earnings and column (3) presents estimates for W-2 earnings; the figures in both of these columns are copied directly from Table 1. In addition, column (2) reports estimates for an earnings series constructed by applying a simple non-parametric correction for censoring to the FICA earnings data. The correction procedure is described in detail in Section 6 of the Appendix. Briefly, data are adjusted for censoring by using the fraction with recorded earnings at the taxable maximum, combined with mean earnings above the taxable maximum estimated from Current Population Surveys, to estimate population mean earnings from censored mean earnings. Note that the effects estimated using the adjusted data are usually bracketed by the effects estimated using the unadjusted FICA and W-2 data. Therefore, only Wald estimates constructed from the adjusted data are reported in the table.

Wald estimates for adjusted FICA earnings, reported in column (5) of Table 3, indicate that white veterans suffered an annual earnings loss of roughly \$2000 constant (1978) dollars of \$3,500 current dollars. This is approximately 15 percent of annual W-2

TABLE 2—VETERAN STATUS AND DRAFT ELIGIBILITY

Data Set	Cohort	Sample	P(Veteran)	$\hat{p}^e$	$\hat{p}^n$	$\hat{p}^e - \hat{p}^n$
Whites						
SIPP (84)[a]	1950	351	0.2673	0.3527	0.1933	0.1594
			(0.0140)	(0.0325)	(0.0233)	(0.0400)
	1951	359	0.1973	0.2831	0.1468	0.1362
			(0.0127)	(0.0390)	(0.0180)	(0.0429)
	1952	336	0.1554	0.2310	0.1257	0.1053
			(0.0114)	(0.0473)	(0.0146)	(0.0495)
	1953	390	0.1298	0.1581	0.1153	0.0427
			(0.0106)	(0.0339)	(0.0152)	(0.0372)
DMDC/CWHS[b]	1950	16119	0.0633	0.0936	0.0279	0.0657
			(0.0019)	(0.0032)	(0.0019)	(0.0037)
	1951	16768	0.1176	0.2071	0.0708	0.1362
			(0.0025)	(0.0053)	(0.0024)	(0.0059)
	1952	17703	0.1515	0.2683	0.1102	0.1581
			(0.0027)	(0.0065)	(0.0027)	(0.0071)
	1953	17749	0.1343	0.1548	0.1268	0.0280
			(0.0026)	(0.0053)	(0.0029)	(0.0060)
Nonwhites						
Data Set	Cohort	Sample	P(Veteran)	$\hat{p}^e$	$\hat{p}^n$	$\hat{p}^e - \hat{p}^n$
SIPP (84)[a]	1950	70	0.1625	0.1957	0.1354	0.0603
			(0.0292)	(0.0699)	(0.0491)	(0.0854)
	1951	63	0.1703	0.2014	0.1514	0.0500
			(0.0292)	(0.0827)	(0.0448)	(0.0940)
	1952	52	0.1332	0.1449	0.1287	0.0161
			(0.0275)	(0.1040)	(0.0373)	(0.1105)
	1953	55	0.1749	0.1980	0.1612	0.0367
			(0.0305)	(0.0865)	(0.0470)	(0.0984)
DMDC/CWHS[b]	1950	5447	0.0417	0.0548	0.0271	0.0276
			(0.0027)	(0.0042)	(0.0032)	(0.0053)
	1951	5258	0.0794	0.1173	0.0599	0.0574
			(0.0037)	(0.0076)	(0.0040)	(0.0086)
	1952	5493	0.0953	0.1439	0.0794	0.0644
			(0.0040)	(0.0095)	(0.0042)	(0.0104)
	1953	5303	0.0925	0.0984	0.0904	0.0080
			(0.0040)	(0.0079)	(0.0046)	(0.0092)

Notes: Standard errors in parentheses. $\hat{p}^e$ is the probability of being a veteran conditional on being draft eligible; $\hat{p}^n$ is the probability of being a veteran conditional on being ineligible.

[a]Wave I, Panel I of the 1984 Survey of Income and Program Participation. Probabilities are for service in the Vietnam era. Estimates are weighted by the SIPP sampling weight and smoothed over 3 cohorts.

[b]Defense Manpower Data Center Administrative Records' information on accessions, from 1970–73, combined with information on cohort size from the Social Security Administration Continuous Work History Sample.

compensation for white men between 1981 and 1984. The similarity of coefficient estimates across cohorts and years suggests that the Wald estimates provide a robust measure of the impact of military service. Taken individually, however, few of the estimates are statistically significant at conventional levels.[7]

[7]The asymptotic standard error of the Wald estimates is derived from the limiting distribution of

TABLE 3—WALD ESTIMATES

		Draft-Eligibility Effects in Current $				
Cohort	Year	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).

B. *Efficient Instrumental Variables Estimates*

The Wald estimator is based solely on earnings differences by draft-eligibility status. A more efficient estimator exploits all the information on RSNs in the aggregate data by fitting earnings model (1) to observations on mean earnings for each group of five consecutive lottery numbers. Consider the following grouped version of (1), where $\bar{y}_{ctj}$ is mean earnings for members of cohort c at time t with lottery numbers in group j, and $\hat{p}_{cj}$ is the fraction of cohort c with lottery numbers in group j who served:

$$\bar{y}_{ctj} = \beta_c + \delta_t + \hat{p}_{cj}\alpha + \bar{u}_{ctj}. \tag{3}$$

Intuitively, estimation of (3) simply generalizes Wald's method to grouped data with more than two groups.

Generalized Least Squares (GLS) estimates of (3) may easily be shown to have an instrumental variables interpretation (Angrist, 1988). In this case, the instrument set includes dummy variables that indicate

$\sqrt{n}(\bar{y}^e - \bar{y}^n)/(\hat{p}^e - \hat{p}^n)$. The standard error is therefore equal to $1/(\hat{p}^e - \hat{p}^n)$ times the standard error of the numerator because the numerator has a nondegenerate limiting distribution, while $(\hat{p}^e - \hat{p}^n)$ converges to a constant. The same standard error formulas arise from application of conventional Instrumental Variables formulas.

groups of five consecutive lottery numbers for each race, cohort, and year of earnings. There are 73 dummy variables for a particular race, cohort, and year; the first indicates men with lottery numbers 1–5 and the 73rd indicates men with lottery numbers 360–365. Furthermore, the quadratic form minimized by the GLS estimator is an overidentification test statistic associated with the use of dummy variables as instruments. This statistic tests the exclusion of lottery number group dummies from equation (1). It may also be viewed as a measure of the goodness-of-fit of the cell means to equation (1).[8]

In principle, implementation of the estimation strategy based on (3) is straightforward—the estimates are simply coefficients from GLS regressions of mean Social Security earnings on estimates of $\hat{p}_{cj}$. The SIPP sample is too small to allow accurate estimation of a full set of $\hat{p}_{cj}$, however. Thus, a second set of probabilities was estimated from a combination of Defense Manpower Data Center (DMDC) administrative records and CWHS data on cohort size. Detailed descriptions of the DMDC administrative records may be found in Angrist (1989b). Briefly, the DMDC data show the total number of new entrants to the military by race, cohort, and lottery number from July 1970 through December 1973.

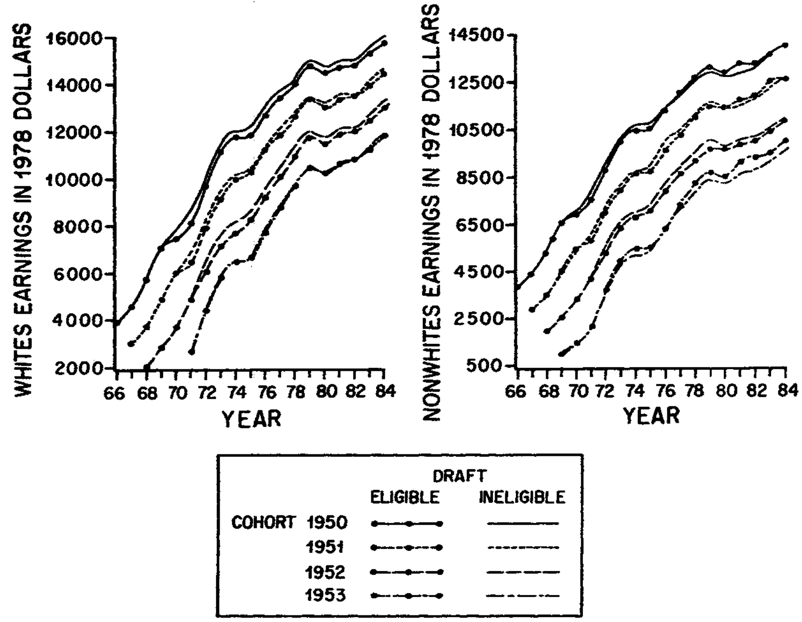
DMDC and CWHS administrative records are used to estimate $\hat{p}_{cj}$ by first counting the number of entrants to the military by race, cohort, and lottery number interval. These numbers are the numerator of the $\hat{p}_{cj}$. Estimates of overall cohort size, to be used in the denominator of $\hat{p}_{cj}$, are derived from the CWHS. Recall that the CWHS is a one percent sample, so that if the CWHS sampling frame is identified with the population at risk, an estimate of total cohort size is simply 100 times the CWHS cohort size. For example, to estimate the probability of being a veteran conditional on being draft eligible, the number of draft-eligible men in the DMDC data is divided by 100 times the number of men in the CWHS with lottery numbers below the induction ceiling. Standard errors for these estimates are computed by applying the usual formula for a binomial proportion.[9]

For comparison with the SIPP estimates, DMDC/CWHS estimates of $\hat{p}^e$ and $\hat{p}^n$ are reported in the lower panel of Table 2. These figures show that, with the exception of the 1950 cohort, the SIPP and DMDC/CWHS procedures give reasonably similar estimates of $\hat{p}^e-\hat{p}^n$. Inaccuracy of the DMDC/CWHS estimates for 1950 is a consequence of the fact that DMDC administrative records are unavailable before July 1970. Therefore, despite the limitations of the SIPP data, the SIPP must be used to construct probabilities for the 1950 cohort. The SIPP sample is too small to allow estimation of a complete set of $\hat{p}_{cj}$ for all lottery number cells in 1950. Consequently, SIPP estimates for 1950 are computed for only two cells, defined by draft-eligibility status, and CWHS earnings data for men born in 1950 are also grouped by draft eligibility. Thus, for each race and year, the sample used to estimate equation (3) includes 73 cell means for each of the three cohorts born from 1951–53, plus two cell means for the 1950 cohort.

A graphical version of equation (3) is depicted in Figure 3, which shows the relationship between probabilities of veteran status ($\hat{p}_{cj}$) and mean W-2 compensation in 1978 dollars ($\bar{y}_{ctj}$) between 1981 and 1984. Plotted in the figure are the average (over four years of earnings) residuals from a regression

[8] A general reference on overidentification testing is Whitney Newey (1985). See also Angrist (1988), where GLS on grouped data is shown to be the minimum variance linear combination of all the Wald estimators that can be computed from any division of grouped observations into linearly independent pairs. The overidentification test statistic for dummy variable instruments is also shown to be the same as the Wald statistic for equality of alternative Wald estimates.

[9] The formula used is $\sqrt{[\hat{p}(1-\hat{p})/n_c]}$, where $\hat{p}$ is the estimated proportion of servers and n_c is the number in the CWHS cohort. For example, 5749 draft eligible white men in the CWHS were born in 1951, and DMDC administrative records show that 119,062 draft-eligible white men born in 1951 served between July 1970 and December 1973. $\hat{p}^e$ is therefore $119{,}062/574{,}900 = 0.21$, with estimated variance equal to $(0.21*0.79)/5749$.



Notes: The figure plots mean W-2 compensation in 1981–84 against probabilities of veteran status by cohort and groups of five consecutive lottery numbers for white men born 1950–53. 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

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

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

of earnings and probabilities on period and cohort effects.[10] Thus, the slope of the ordinary least squares regression line drawn through the points corresponds to an estimate of α. This slope is equal to $-2{,}384$ dollars, with a standard error of 778 dollars. An interesting feature of the figure is the apparent heteroscedasticity of the earnings residuals. Dispersion around the regression line is reduced for cells with high probabilities of veteran status. This heteroscedasticity also appears in comparisons (not shown here) of cell variances by draft-eligibility status; draft-eligible men have somewhat less variable earnings.

As pointed out earlier, estimation of (3) is the same as instrumental variables estimation of (1) using dummy variables as instruments. However, inference for the case where the estimation strategy is implemented by regressing CWHS mean earnings on DMDC/CWHS or SIPP probabilities is complicated by the use of data from multiple samples. Assuming that the samples used to calculate mean earnings and the sample used to calculate $\hat{p}_{cj}$ are independent, the optimal Two-Sample Instrumental Variables (TSIV)

[10]There are 221 points plotted in the figure: 4 years of earnings times 3 cohorts with 73 cells plus 4 years of earnings times one cohort with 2 cells (men born in 1950) = 884, divided by four to compute the average over years.

estimator has a simple form that may be briefly described as follows. Let $\bar{y}$ denote the vector of $\bar{y}_{ctj}$, $\hat{p}$ denote the vector of $\hat{p}_{cj}$, and $\bar{u}(\theta)$ denote the vector of $\bar{u}_{ctj}$, where θ in parentheses represents the dependence of residuals on the parameter vector. Also, let $V(\ )$ denote the covariance matrix of the argument. Then the optimal TSIV estimator chooses θ to minimize

$$\bar{u}(\theta)'\left[V(\bar{y})+\alpha^2V(\hat{p})\right]^{-1}\bar{u}(\theta)\equiv m(\theta),$$

which is also the GLS minimand for (3).[11]

The minimized value of $m(\theta)$ is an overidentification test statistic for the validity of dummy variables as instruments. If some of these dummy variables are correlated with the regression error, then $m(\theta)$ should be large relative to a chi-square distribution with degrees of freedom equal to the difference between the number of instruments and the number of estimated parameters.

Table 4 presents two sets of TSIV estimates of equation (3) for 1981–84 earnings in 1978 dollars. Model 1 allows the effect of veteran status on earnings to vary by cohort, while Model 2 restricts estimated service effects to be the same across cohorts. Note that, as in Table 3, the heading "adjusted FICA earnings" refers to FICA-taxable earnings adjusted for censoring at the taxable maximum using the procedure described in section 6 of the Appendix.

The results in Table 4 show that white veterans born from 1950–52 suffered an annual earnings loss of between \$1,500 and \$2,100 constant (1978) dollars. These results generally are similar in magnitude to the Wald estimates reported in Table 3. Also, as in Table 3, regression estimates for adjusted FICA earnings tend to be bracketed by the results for unadjusted FICA and W-2 earnings. Although many of the estimates for individual cohorts in Model 1 are not significant, the combined estimates for whites in Model 2 are substantially larger than twice their standard errors. In contrast, results for nonwhites show no evidence of a statistically significant earnings loss to veterans.

The overidentification test statistics reported in Table 4 take on values less than their degrees of freedom, suggesting that the residuals in equation (1) are not correlated with lottery-based instruments.[12] It should be noted, however, that low values of the test statistics may indicate low power in a test with so many degrees of freedom. On the other hand, without a particular alternative hypothesis in mind, it seems natural to report the omnibus goodness-of-fit test.

Subtracting the test statistic for Model 1 from the test statistic for Model 2 gives a chi-square test for the restriction of equal treatment effects across cohorts. The set of restrictions imposed by equal treatment effects has three degrees of freedom. None of the chi-square statistics for Model 2 are larger than the corresponding statistics for Model 1 by as much as three, indicating that

[11]See Angrist (1989d) for details. Estimates of α for use in the formula for Φ were computed by weighted least squares using the inverse of the sampling variance of the $\bar{y}_{ctj}$ as weights. Estimates of $V(\bar{y})$ and $V(\hat{p})$ are discussed in Section 5 of the Appendix. Note that the TSIV estimator may also be motivated as an application of Optimal Minimum Distance (OMD) techniques such as those described by Gary Chamberlain (1982). Ignoring period and cohort effects, OMD estimates for the current problem are tabulated by choosing α and p to minimize

$$q(\alpha,p)=$$

$$\begin{bmatrix}\bar{y}-p\alpha\\ \hat{p}-p\end{bmatrix}'\begin{bmatrix}V(\bar{y}) & 0\\ 0 & V(\hat{p})\end{bmatrix}^{-1}\begin{bmatrix}\bar{y}-p\alpha\\ \hat{p}-p\end{bmatrix}.$$

By concentrating out the estimate of p, it is easy to show that $q(\alpha,p)=m(\theta)$.

[12]Degrees of freedom for the overidentification tests are calculated as follows. For each race, the data consist of four years of earnings for three cohorts with 73 lottery number cells each. The fourth cohort, men born in 1950, has four years of earnings with 2 lottery number cells each. This gives a total of 884 cells or, equivalently, 884 categorical instruments. Model 1 includes 4 cohort dummies, 3 year dummies, and 4 treatment effects. 884 minus 11 parameters gives 873 degrees of freedom. Model 2 has 3 fewer parameters than model 1 and consequently the chi-square statistic for model 2 has 876 degrees of freedom.

TABLE 4—TWO-STAGE INSTRUMENTAL VARIABLES ESTIMATES

Whites			
Cohort	FICA Taxable Earnings	Adjusted FICA Earnings	Total W-2 Compensation
Model 1			
1950	−1709.2	−2093.7	−1895.0
	(946.8)	(1108.8)	(1333.1)
1951	−1457.1	−1983.7	−2431.4
	(959.3)	(1036.1)	(1152.1)
1952	−1724.0	−1943.0	−2058.7
	(863.1)	(927.2)	(1001.9)
1953	1223.8	900.7	−488.6
	(3232.1)	(3505.3)	(3936.0)
$\chi^2(873)$	578.3	630.3	569.5
Model 2			
1950–53	−1562.9	−1920.4	−2094.5
	(521.8)	(575.9)	(646.3)
$\chi^2(876)$	579.1	631.0	569.7

Nonwhites			
Cohort	FICA Taxable Earnings	Adjusted FICA Earnings	Total W-2 Compensation
Model 1			
1950	3893.7	3891.9	5711.8
	(5358.5)	(6244.5)	(7206.0)
1951	−891.3	−333.4	2609.0
	(4397.1)	(4664.2)	(4894.6)
1952	−3182.9	−3457.7	−3068.0
	(3997.4)	(4195.2)	(4229.2)
1953	−5928.3	−8571.4	−6325.8
	(10296.3)	(10697.1)	(11410.6)
$\chi^2(873)$	616.7	681.7	693.6
Model 2			
1950–53	−643.3	−999.7	366.7
	(2407.5)	(2602.5)	(2734.2)
$\chi^2(876)$	618.4	683.4	695.6

Notes: Standard errors in parentheses.
The table shows estimates of the effect of military service on average 1981–84 earnings in 1978 dollars. The estimation method is optimally weighted Two-Sample Instrumental Variables, described in the text. FICA and W-2 earnings are from the Social Security CWHS. The adjusted FICA series is described in the Appendix.

the estimated treatment effects are not statistically different across cohorts.

IV. Military Service and Loss of Labor Market Experience

The simplest explanation for a veteran earnings penalty is that military experience is a poor substitute for lost civilian labor market experience. As evidence for this hypothesis, Griliches and Mason (1972) show that the longer they were in the military, the less veterans earn relative to nonveterans. A test of the loss-of-experience hypothesis is developed here using the functional form commonly employed in empirical studies of 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.

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).

experience. Excluding the time-varying intercept, model (6) contains four structural parameters; β_0, β_1, γ, and l, and four reduced form parameters; β_0, γ, π_1 and π_2. Model (5) imposes one testable restriction on the reduced form by setting $\beta_1 = 0$.[15]

Table 5 shows results from nonlinear GLS estimation of (5) and (6), and results from Linear GLS estimation of (7), using data on the real FICA earnings of white men born from 1950 to 1952. The weighting matrix used in estimation was derived in a manner similar to the weighting matrix used to construct the estimates in Table 4.[16] Because earnings functions are commonly fit in logs, the dependent variable is taken to be the log of mean earnings for each cell. The log of the mean is not the same as the mean of the log, but the CWHS data set does not contain the mean of log earnings. If earnings are approximately lognormally distributed, use of the log of the mean will provide a reasonable approximation. In practice, estimates of

[15] A third model is derived by letting both γ and β vary with veteran status. This model leads to a reduced form similar to (7), with the only modification being the addition of a linear term of the form $\pi_3(\hat{p}_{cj} * x_{ct}^2)$. In the empirical work, however, no evidence was found that such a term belongs in the earnings function reduced form.

[16] The only modification arises from the fact that for equations (5)–(7), reduced form treatment effects appearing in the weighting matrix are time varying. Let Π_c denote the vector of α_{ct} for the time-series of earnings by cohort c and suppose that each time-series is of length T. Then the second term in the optimal weighting matrix has the following block corresponding to the time series of earnings for lottery number cell j of cohort c:

$$(1/n_{cj})\Pi_c \times \{[e_T e_{T'}] \otimes [\hat{p}_{cj}(1-\hat{p}_{cj})]\}\Pi_c',$$

where e_T is a vector of T 1's. In practice, Π_c is replaced by weighted least squares estimates (weights are the inverse sampling variances of $\bar{y}_{ctj}$) of the reduced form equation, (7). Estimates of models for the log of earnings replace $V(\bar{y})$ with delta-method estimates of the covariance matrix of $\log(\bar{y})$.

models in levels resulted in inferences similar to those arising from estimates of models in logs.

The sample used to estimate equations (5)–(7) begins in the fifth year after the lottery in which members of the cohort were drafted. For example, the sample begins in 1975 for men born in 1950. This allows for three years of service and one year of readjustment to civilian life. Median length of service of Vietnam era veterans was 37 months (Veterans Administration 1981b, p. 16). Evidence from Table 4 suggests that veteran effects are essentially zero for the 1953 cohort, so it was excluded from the estimation.

Estimates of models (5), (6), and (7) are presented in columns (1), (2), and (3) of Table 5. The chi-square statistic in column (3) is an overidentification test statistic for the overall goodness-of-fit of the reduced form and the chi-square statistic in column (1) is for the restriction $\beta_1 = 0$. Restricting β_1 to be zero does not affect the overall fit. The effect of veteran status on earnings growth, although negative, is not statistically significant. The sum of the test statistics in columns (1) and (3) is an overidentification test statistic for the overall goodness-of-fit of model (5). This statistic also takes on a value less than its degrees of freedom, indicating that the simplest loss-of-experience model is not at odds with the data.

The loss of experience estimated using model (5) is approximately 2 years. This is somewhat low relative to the median length of service, suggesting that military experience may be a partial substitute for civilian experience. Table 5 also shows the age at which the reduced form veteran effect finally reaches zero. This occurs when $x_{ct} = -\pi_1/\pi_2$, so that $\alpha_{ct} = 0$. The reduced form estimates in column (3) imply that the loss of earnings to veterans decays to zero around age 50 with a standard error[17] of 16.

[17]Delta method standard errors are given by the square root of $(\sigma_1/\pi_2^2) - (2\sigma_{12}\pi_1/\pi_2^3) + (\sigma_2\pi_1^2/\pi_2^4)$, where σ_1, σ_2, and σ_{12} are the elements of the covariance matrix of the estimated π_1 and π_2. σ_{12} is estimated to be -0.00019, and the square roots of σ_1 and σ_2 appear in Table 5.

V. Caveats

The consistency of lottery-based estimates of the effects of military service turns on two key assumptions. First, earnings model (1) must be an accurate representation of the impact of military service. Second, functions of the draft lottery must be valid instruments for s_i in the earnings regressions. Three models leading to the failure of these assumptions are discussed briefly below. The first two, incorporating treatment effect heterogeneity and missing covariates, merely result in a reinterpretation of the estimates. The third model, allowing for earnings-modifying draft-avoidance behavior, calls into question the assumption that earnings differences by lottery number can be attributed solely to veteran status.

A. *Treatment Effect Heterogeneity*

The effect of veteran status can only be estimated for those who served in the armed forces. If veterans are more or less likely to benefit from military service than the rest of the population, then the estimated veteran status coefficient does not characterize the impact of veteran status on a random sample. This problem is formalized in the random coefficients model for treatment effect heterogeneity:

$$y_{cti} = \beta_c + \delta_t + s_i\alpha_i + u_{it}$$
$$\alpha_i = \alpha_0 + \varepsilon_i,$$

where α_i is the effect of military service on the earnings of person i, with population mean equal to α_0. If ε_i is uncorrelated with the instruments, an instrumental variables estimator of the random coefficients model identifies $\alpha_0 + E(\varepsilon_i | s_i = 1)$, and not α_0 (James Heckman and Richard Robb, 1985).

Related to the problem of treatment effect heterogeneity is the fact that not all Vietnam era accessions to the military were induced by the draft. A substantial fraction of enlistments were made by "true volunteers" who would have volunteered for service in the absence of a draft (Tarr, 1981). Suppose that the impact of military service on the earnings of true volunteers differs from the impact on the earnings of draftees and men who enlisted because of the draft. Then using functions of the draft lottery as instruments will only identify the effect of military service for the latter group. To see this, let f_i be an indicator for the veteran status of draftees and draft-motivated volunteers and let g_i be an indicator of veteran status for true volunteers. Write

$$\alpha_i s_i = \alpha_f f_i + \alpha_g g_i$$

for the treatment effect experienced by i. Note that functions of draft lottery numbers will only be correlated with f_i. Consequently, α_f is identified by lottery-based IV estimators, but $\alpha_g g_i$ becomes part of the regression error.

B. *The Absence of Covariates*

An additional problem arises because Social Security data contain no information on covariates other than race and age. This may be of concern if the impact of veteran status is primarily through its effect on covariates. For example, veterans might have a higher level of educational attainment because of financial aid available through the GI Bill. In the absence of data on education, estimated veteran effects confound the "pure" effect of military service with its effect on education. Formally, the need for covariates may be represented by replacing δ_t with $x_i\delta$ in (1). In this case, instrumental variables estimates identify

$$\bar{\alpha} = \left[E(x_i|s_i=1) - E(x_i|s_i=0)\right]\delta + \alpha.$$

In many applications, however, it may be $\bar{\alpha}$ that is actually the parameter of interest. For example, the fact that veteran status influences civilian earnings primarily through its influence on third variables might be of little importance for issues related to veterans compensation.

C. *Earnings-Modifying Draft Avoidance Behavior*

Perhaps the most serious problem arises if the risk of induction affected earnings for some other reason than through an effect on the probability of military service. For example, it is sometimes argued that during the Vietnam era, students went to college to avoid the draft and that educational standards were reduced to avoid having to flunk students out of school. Lawrence Baskir and William Strauss (1978) claim that Vietnam era college enrollment was 6–7 percent higher than normal because of the draft.

If draft-avoidance behavior is correlated with lottery numbers and with variables related to earnings besides veteran status, then lottery-based instruments will be correlated with the regression error in equation (1). Such correlation will bias estimates of the effects of military service constructed using the lottery. But in a previous study using micro data (Angrist, 1989a), specification tests provided no evidence of a relationship between lottery numbers and characteristics other than veteran status. The overidentification test statistics reported here also show no evidence of omitted variables bias. Finally, even if having a low lottery number is correlated with a tendency to stay in school, the fact that earnings rise with schooling implies that lottery-based estimates of the effect of veteran status will tend to underestimate the true effect.

VI. Conclusions

Estimates based on the draft lottery indicate that as much as ten years after their discharge from service, white veterans who served at the close of the Vietnam era earned substantially less than nonveterans. The annual earnings loss to white veterans is on the order of \$3,500 current dollars, or roughly 15 percent of yearly wage and salary earnings in the early 1980s. In contrast, the estimated veteran effects for nonwhites are not statistically significant.

In light of the results reported here, some more conventional estimates of the effect of Vietnam era veteran status do not appear to be too far off the mark. Rosen and Taubman (1982) report estimates close to these, finding a 19 percent annual earnings loss to Vietnam era veterans. Crane and Wise (1987) find an 11 percent reduction in 1979 weekly earnings. On the other hand, Mark Berger and Barry Hirsch (1983) find essentially no effect on 1977 weekly earnings, and Angrist (1989a) reports OLS coefficients of zero when using the National Longitudinal Survey (NLS) to estimate the effects of veteran status on hourly wages. In contrast to the OLS estimates, lottery-based estimates from the NLS indicate that white veterans had lower hourly wages in 1981 than their nonveteran counterparts. Similar results from the SIPP are reported in Angrist (1989c). Thus, lottery-based estimates from a variety of sources provide conclusive evidence that white Vietnam veterans were disadvantaged by their service.[18]

This paper also proposes a simple explanation for the loss of earnings to white veterans: they earn less because their military experience is only a partial substitute for the civilian labor market experience lost while in the armed forces. Goodness-of-fit tests suggest that for whites, the time-series of veteran status coefficients is consistent with this hypothesis. Experience-earnings profiles estimated using Social Security data imply that white veterans suffered an earnings reduction equivalent to the loss of two years civilian labor market experience.

The analysis reported here leads naturally to further research on a number of topics. One of these is the question of alternatives to the loss-of-experience explanation for the reduction in white veterans' earnings. Veteran status may be a screening device, as suggested by DeTray (1982), or there may be cohort size effects such as those discussed by Finis Welch (1979). Because the Social Security data include information on variances, testable implications of these theories might also include restrictions on second, as well as first, moments. Another question for future research is whether draft eligibility affected educational and career plans independently of its effect on military service. The lottery may provide a useful tool for research on changing educational attainment in the 1960s and 1970s.

Finally, there remains the question of reconciling the loss of earnings to Vietnam era veterans with the apparent benefits of military service to veterans of World War II and other eras (Rosen and Taubman, 1982; Berger and Hirsch, 1983). Elsewhere, Alan Krueger and I have argued that the need for reconciliation is, at least in part, illusory (Angrist and Krueger, 1989). Although OLS regressions usually show that the effect of World War II veteran status is large, positive, and significant, these results may actually be a consequence of selection bias. By exploiting the fact that World War II veteran status is also correlated with exact date of birth, we have implemented an instrumental variables estimation strategy similar in spirit to the one used here. The results of this procedure indicate that the true impact of World War II veteran status on earnings is no larger than zero and may well be negative.

APPENDIX: DATA SOURCES AND METHODS

1. CWHS Data Collection[19]

The Social Security Administration maintains the earnings histories of covered employees in a data base known as the Summary Earnings Record (SER). Approximately one year after the SER has been updated with the latest year's earnings, a one percent sample of earnings histories is extracted. The sampling frame consists of all issued Social Security numbers, and the sample is stratified using the regional information coded in the numbers.

Prior to 1978, the FICA taxable earnings of employees were reported to the SSA by employers on a quarterly basis. Self-employed workers report their earnings annually on schedule SE of Internal Revenue Service (IRS) Form 1040, which is forwarded to the SSA by the IRS. Since 1978, employers have no longer been required to make quarterly reports. Instead they file IRS form W-2 with the SSA, as well as with the IRS, on an annual basis.

After 1978, all earnings, including those above the FICA taxable maximum, are to be reported to the SSA on form W-2. Furthermore, all employers are required

[18] Lottery-based estimates from the SIPP and NLS suggest that nonwhite veterans may have higher monthly and hourly earnings than comparable nonveterans. The estimates for nonwhites are too imprecise, however, to be viewed as conclusive.

[19] This section draws on Warren Buckler and Cresston Smith (1984), U.S. Department of Health and Human Services (1987), and personal correspondence and conversations with Buckler and Smith.

to file W-2s with the SSA regardless of whether their employees are engaged in FICA taxable employment. In practice, however, many employers do not report the earnings of those engaged in non-FICA-taxable employment. A further shortcoming of the W-2 series is the poor quality of the data during the first years of annual reporting.

2. Coverage and Truncation of the CWHS Earnings Series

FICA coverage includes most wage and salary and self-employment earnings. For the sample period used here, the most important coverage exceptions are the majority of federal civilian employees, some state and local government employees, some agricultural and domestic workers, and the employees of some nonprofit organizations. A view of coverage by industry in 1981 is given in Robert Meyer (1985, Table 2.1). Meyer's figures for state, local, and civilian federal government employees imply that by 1981, roughly 58 percent of all civilian government workers were covered.

FICA taxable maximums are reported in Appendix Table A1. The combined effects of limited coverage and censoring at the taxable maximum are conveniently summarized by the percentage of all earnings that are reported to the SSA. These statistics, reported in Department of Health and Human Services (1987, Table 30), show that after 1981 over 90 percent of wage and salary earnings and over 75 percent of self-employment earnings were reported to the SSA.

The W-2 earnings series excludes earnings from self-employment. Unpublished estimates indicate that roughly 6.4 percent of men born between 1944 and 1953 with nonzero earnings in 1984 had self-employment earnings only (figures from private correspondence with SSA employees). Other differences between FICA and W-2 earnings coverage are described by Mary Millea and Beth Kilss (1980).

3. Matching Dates of Birth to the CWHS

In the CWHS, information on race and sex is obtained from a computerized record of applications for a Social Security number called the NUMIDENT file. For this project, dates of birth were also matched from the NUMIDENT file to the CWHS. Draft lottery numbers were then matched to the dates of birth using lottery number tables in Selective Service System Semiannual Reports for 1969–73. A small number of individuals were discarded from the final data set because there was no information on either their sex, race, or exact date of birth.

4. CWHS Descriptive Statistics

Descriptive statistics for the CWHS cohorts studied here are reported in Table A1. The combined statistics for four cohorts were constructed by computing weighted averages of cohort means. Unless otherwise noted, statistics in the table refer to men with positive earnings. Sample sizes decline over time due to attrition from mortality.

The descriptive statistics indicate that after 1972 the fraction of men in the sample with zero FICA earnings varies roughly between 15 and 22 percent for whites and between 33 and 36 percent for blacks. To evaluate these figures, note that the author's tabulations show that roughly 10 percent of white men in these cohorts have zero recorded wage and salary earnings in the late 1970s' Current Population Surveys. Suppose that of the 90 percent who work, 12 percent are in the uncovered sector so that only 79 percent of the cohort may be expected to have positive FICA wage and salary earnings. Adding an estimated 5 percent who only have FICA self-employment earnings implies that 16 percent should have zero FICA earnings of any type. Thus, 14–18 percent of CWHS earnings being zero for whites between 1973 and 1980 does not seem unreasonable.

In recent years, the fraction with zero earnings appears to be too high to be accounted for by labor force participation or employment in the uncovered sector. This is probably because of the long delay in filing and recording Social Security taxable earnings. In their analysis of Social Security data, Card and Sullivan (1988) also note the problems caused by filing delay. Problems of undercoverage and filing delay may be especially severe for nonwhites. The 15 percent of CPS nonwhites with zero wage and salary earnings is not large enough to explain the approximately 34 percent of nonwhites with zero FICA earnings in the CWHS.

The fraction with FICA earnings at the taxable maximum is more variable than the fraction with zero earnings, ranging from 3 to 15 percent for whites and between 2 to 10 percent for blacks. The FICA earnings of men with multiple employers can exceed the taxable maximum because reported earnings are censored by source. The fraction of men with FICA earnings above the taxable maximum is around 1–2 percent. The W-2 and FICA earnings series show roughly equal fractions at or above the FICA taxable maximum, suggesting that both variables are drawn from the same underlying distribution. But problems with early years of the W-2 series are clearly reflected in the sample statistics. For example, the standard deviation of whites' W-2 earnings in 1978 is six times the mean and does not fall below the mean until 1981. Another disturbing feature of the W-2 series is that *nominal* earnings appear to fall from 1978–80. The W-2 series also has a substantially higher fraction of zeros than the FICA series does. However, part of this difference is caused by the inclusion of self-employment earnings in the FICA series. Also, generally there are some individuals with FICA taxable earnings but no federally taxable compensation (Millea and Kilss, 1980).

5. Covariance Estimates for Social Security Earnings and for $\hat{p}_{cj}$

Information on second moments in the aggregated CWHS is restricted to variances. Therefore, off-diagonal elements of the covariance matrix of Social Security earnings must be estimated. Recall that $\bar{y}$ denotes the vector of $\bar{y}_{ctj}$, where c indexes cohort, t indexes year of earnings, and j indexes lottery number cells. The covariance matrix of $\bar{y}$ is block diagonal, with nonzero elements for the covariance between $\bar{y}_{ctj}$ and $\bar{y}_{ckj}$, and zeros everywhere else. Thus, there is only correlation between elements of the time-series of earnings for a particular cohort and lottery number group. The in-

TABLE A1—DESCRIPTIVE STATISTICS FOR MEN BORN 1950–53

Race	Year	N	Taxable Maximum	FICA Earnings	FICA at Limit[a,b]	FICA Zeros	W-2 Earnings[c]	W-2 at Limit[a]	W-2 Zeros[a]
White	69	68,407	7,800	1,473 (1,457)	0.003	0.321			
	70	68,339	7,800	1,977 (1,825)	0.010	0.241			
	71	68,244	7,800	2,581 (2,233)	0.030	0.194			
	72	68,154	9,000	3,614 (2,802)	0.040	0.157			
	73	68,053	10,800	4,738 (3,460)	0.046	0.141			
	74	67,966	13,200	5,727 (4,170)	0.034	0.145			
	75	67,882	14,100	6,459 (4,843)	0.053	0.169			
	76	67,794	15,300	7,698 (5,548)	0.078	0.167			
	77	67,691	16,500	8,974 (6,206)	0.107	0.163			
	78	67,598	17,700	10,441 (7,050)	0.149	0.165	15,435 (91,032)	0.167	0.244
	79	67,503	22,900	12,388 (8,455)	0.092	0.166	14,786 (65,359)	0.097	0.261
	80	67,413	25,900	13,769 (9,678)	0.092	0.175	14,561 (28,180)	0.096	0.272
	81	67,316	29,700	15,641 (11,129)	0.089	0.183	16,363 (15,295)	0.092	0.272
	82	67,265	32,400	16,743 (12,371)	0.092	0.200	17,907 (16,298)	0.093	0.282
	83	67,190	35,700	18,046 (13,621)	0.089	0.210	19,595 (22,470)	0.089	0.275
	84	67,114	37,800	19,717 (14,883)	0.103	0.219	21,595 (20,856)	0.101	0.281
Nonwhite	69	21,514	7,800	1,306 (1,471)	0.002	0.464			
	70	21,501	7,800	1,711 (1,865)	0.007	0.417			
	71	21,475	7,800	2,197 (2,275)	0.020	0.388			
	72	21,443	9,000	3,072 (2,922)	0.025	0.352			
	73	21,412	10,800	3,992 (3,609)	0.029	0.330			
	74	21,380	13,200	4,802 (4,359)	0.021	0.336			
	75	21,339	14,100	5,404 (5,015)	0.031	0.359			
	76	21,310	15,300	6,453 (5,836)	0.049	0.350			
	77	21,275	16,500	7,510 (6,616)	0.071	0.341			
	78	21,243	17,700	8,751 (7,541)	0.096	0.338	13,439 (63,722)	0.116	0.391
	79	21,200	22,900	10,262 (8,938)	0.061	0.329	13,581 (63,244)	0.067	0.393
	80	21,167	25,900	11,405 (10,147)	0.063	0.331	11,716 (13,895)	0.067	0.397
	81	21,130	29,700	12,986 (11,628)	0.063	0.335	13,421 (14,644)	0.063	0.388
	82	21,109	32,400	14,045 (12,849)	0.067	0.354	14,983 (16,112)	0.065	0.408
	83	21,077	35,700	15,101 (14,167)	0.067	0.358	16,271 (19,089)	0.065	0.396
	84	21,042	37,800	16,391 (15,237)	0.077	0.353	17,905 (20,784)	0.073	0.387

Notes: Statistics are from the Social Security Administration CWHS. Standard deviations of earnings in parentheses. Amounts are in current dollars. Sample statistics are weighted averages of cells for each race, year of birth, and five consecutive lottery numbers.

[a] Fractions at limit are fractions of nonzero earnings at or above FICA taxable maximum. Fractions zero are fractions of all nondecedents with zero earnings.

[b] FICA earnings are wage and salary and self-employment earnings in Social Security taxable employment. FICA taxable earnings are censored at the taxable maximum except for those with multiple sources. Multiple sources are censored by source.

[c] W-2 earnings are total W-2 form wage and salary compensation, not censored at the Social Security taxable maximum. W-2 earnings do not include earnings from self-employment.

tertemporal correlation structure is assumed constant across lottery number groups so that the 73 cells available for each race and cohort may be used to estimate correlation matrices for all race-cohort combinations. Correlations are converted to covariances using the within cell variances available from the CWHS data. This procedure is also applied to the adjusted FICA series described in Section 6 of the Appendix, with the modification that adjusted standard errors (diagonal elements of Ω) are used to convert correlations to covariances.

The covariance matrix of $\hat{p}$ is also block diagonal, with elements equal to the variance of $\hat{p}_{cj}$ in every element of the block corresponding to the time series for cohort c and lottery cell j. The variance of $\hat{p}_{cj}$ is estimated using the standard formula for an estimated proportion. The sample size in the formula is taken to be the size of the SIPP cohort for those born in 1950 and the size of the CWHS cohort for those born from 1951 to 53.

6. FICA Earnings Adjusted for Censoring

For economy of notation, in this section all cells are indexed by j. The relationship between the expectation of censored earnings and the expectation of uncensored earnings for cell j is given by

$$\mu_j^0=\mu_j^c+p_j^l\left(\mu_j^l-L_j\right),$$

where μ_j^0 is the mean of uncensored earnings, μ_j^c is the mean of censored earnings, μ_j^l is the mean of earnings above the taxable maximum, L_j is the taxable maximum, and p_j^l is the fraction with earnings at or above the taxable maximum.[20]

The FICA earnings series is adjusted for censoring by applying this formula using estimates of μ_j^l tabulated from March Current Population Surveys (CPS) for each year, race, and cohort. Although this procedure involves no parametric distributional assumptions, the adjustment is only approximate because the CPS estimates of μ_j^l do not vary by lottery number. However, the adjustment does incorporate variation in p_j^l by lottery number.

Data on earnings above the FICA taxable maximum (μ_j^l) are taken from the Mare-Winship March CPS Uniform files. Cohort was determined on the basis of age in 1985; ages 34–35 were assigned 1950, ages 33–34 were assigned 1951, ages 32–33 were assigned 1952, and ages 31–32 were assigned 1953. The CPS reports wage and salary earnings in the year preceding the survey year. To compute the adjusted 1981–84 earnings series used in Tables 3 and 4, data from the CPSs for 1982–85 were used to compute separate a μ_j^l by cohort and race for each year. Additional details and summary statistics for the CPS data are reported in Appendix A of Angrist (1989c). It should be noted that CPS earnings data are also censored—at \$75,000 for 1981–83 earnings and at \$100,000 for 1984 earnings.

Standard errors for the adjusted series are calculated as follows. Let $m=[m_1'\,m_2'\,m_3']'$ denote the vector of sample moments corresponding to μ^c, p^l, and μ^l and let $\Sigma_{ij}(i,j=1,2,3)$ denote the corresponding blocks of the covariance matrix of m. The covariance matrix of m is assumed to be given by

$$\begin{pmatrix}\Sigma_{11} & \Sigma_{12} & 0\\ \Sigma_{12} & \Sigma_{22} & 0\\ 0 & 0 & \Sigma_{33}\end{pmatrix}.$$

The delta-method covariance matrix for the vector of adjusted earnings, $m_1+m_2(m_3-L)$, is

$$\Omega=\Sigma_{11}+2\Sigma_{12}(m_3-L)$$
$$+\Sigma_{22}(m_3-L)^2+m_2^2\Sigma_{33}.$$

Only the diagonal elements of Ω are estimated using the above formula. Estimates of diagonal elements of Σ_{11} and Σ_{22} are available from the CWHS cell statistics, while diagonal elements of Σ_{33} are estimated from the CPS micro data. An estimator for diagonal elements of Σ_{12} is easily shown to be (Angrist, 1989c)

$$\sigma_{12j}=\left[p_j^l\left(L_j-\mu_j^c\right)\right]/n_j.$$

A simplified procedure, described in Section 5 of this Appendix, is used to estimate the nonzero off-diagonal elements of Ω directly from the adjusted cell means.

7. Matching Draft-Eligibility Status to the SIPP

The Survey of Income and Program Participation is a Census Bureau longitudinal survey of approximately 20,000 households in the civilian noninstitutional population (U.S. Department of Commerce, 1985). Data for the first wave of the first SIPP panel were collected from four rotation groups in 1983 and 1984.

The SIPP public-use tapes contain year and month of birth. A Census Bureau "in-house" version of the

[20] This formula ignores the fact that the censored earnings of men with multiple employers may be above the taxable maximum. The formula may be used to analyze the bias in treatment effects estimated from censored data. Suppose that cell j is for a sample of draft-eligible men and that cell k is for a sample of draft-ineligible men. The draft-eligibility treatment effect estimated from the difference between CWHS censored mean earnings in cells j and k is the sample analogue of

$$\mu_j^c-\mu_k^c=\left(\mu_j^0-\mu_k^0\right)+\left(p_j^l-p_k^l\right)\left(L_j-\mu_j^l\right)$$
$$+p_k^l\left(\mu_k^l-\mu_j^l\right).$$

Assuming that the effect of draft eligibility on both p^l and μ^l is of the same sign as the effect on μ^0, this expression shows that treatment effects estimated from the censored data differ from the true treatment effect by terms that are opposite in sign from the true effect.

SIPP contains information on day of birth, which is not released to the public. At the author's request, this information was used to match a dummy variable for draft-eligibility status to the public use version of SIPP Panel I. Draft eligibility was determined by the official RSN ceiling for men born from 1944–52 and by RSN 95 for men born in 1953. Vietnam era veteran status is coded from the SIPP variables VETSTAT, which records veteran status, and from U-SRVDTE, which records the period of service.

REFERENCES

Angrist, Joshua D., "Grouped Data Estimation and Testing in Simple Labor Supply Models," Industrial Relations Section Working Paper, no. 234, Princeton University, July 1988.

______, (1989a), "Using the Draft Lottery to Measure the Effects of Military Service on Civilian Labor Market Outcomes," in *Research in Labor Economics*, *10*, Ron Ehrenberg, ed., Greenwich: JAI Press, 1989.

______, (1989b), "Selection for Military Service in the Vietnam Era," Industrial Relations Section Working Paper, no. 250, Princeton University, April 1989.

______, (1989c), "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," Industrial Relations Section Working Paper, no. 251, Princeton University, April 1989.

______, (1989d), "Instrumental Variables Regression with Moments from Two-Samples," Chapter 3 in *Econometric Analysis of the Vietnam Era Draft Lottery*, unpublished Princeton University doctoral dissertation, October 1989.

______ **and Krueger, Alan,** "Why Do World War II Veterans Earn More?" National Bureau of Economic Research Working Paper, no. 2991, May 1989.

Baskir, Lawrence M. and Strauss, William A., *Chance and Circumstance: The Draft, The War and the Vietnam Generation*, New York: Alfred A. Knopf, 1978.

Berger, Mark C. and Hirsch, Barry T., "The Civilian Earnings Experience of Vietnam-Era Veterans," *Journal of Human Resources*, Fall 1983, *18*, 455–79.

Buckler, Warren and Smith, Cresston, "The Continuous Work History Sample (CWHS): Description and Contents," in *Statistical Use of Administrative Records: Recent Research and Present Prospects Volume I*, Department of the Treasury, Internal Revenue Service, Statistics of Income Division, March 1984.

Card, David and Sullivan, Daniel, "Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment," *Econometrica*, January 1988, *56*, 497–530.

Chamberlain, Gary, "Multivariate Regression Models for Panel Data," *Journal of Econometrics*, January 1982, *18*, 5–46.

Crane, Jon R. and Wise, David A., "Military Service and the Civilian Earnings of Youths," Chapter 6 in *Public Sector Payrolls*, D. A. Wise, ed., Chicago: University of Chicago Press, 1987.

DeTray, Dennis N., "Veteran Status as a Screening Device," *American Economic Review*, March 1982, *72*, 133–42.

Griliches, Z. and Mason, William, "Education, Income and Ability," *Journal of Political Economy*, May/June 1972, *80*, Part 2: S74–S103.

Hearst, Norman, Newman, Tom B. and Hulley, Stephen B., "Delayed Effects of the Military Draft on Mortality: A Randomized Natural Experiment," *New England Journal of Medicine*, March 6, 1986, *314*, 620–24.

Heckman, James J. and Robb, Richard, "Alternative Methods for Evaluating the Impact of Interventions," in J. Heckman and B. Singer, eds., *Longitudinal Analysis of Labor Market Data*, New York: Cambridge University Press, 1985.

Meyer, Robert J., *Social Security*, 3rd ed., Homewood, IL: Richard D. Irwin, 1985.

Millea, Mary T. and Kilss, Beth, "Exploration of Differences Between Linked Social Security and Internal Revenue Service Wage Data for 1972," in *Studies from Interagency Data Linkages*, Report No. 11: Measuring the Impact on Family and Personal Income Statistics of Reporting Differences Between the Current Population and Administrative Sources, U.S. Department of Health, Education and Welfare, Social Security Administration, March 1980.

Mincer, Jacob, *Schooling, Experience and Earnings*, New York: National Bureau of Economic Research, 1974.

Newey, Whitney K., "Generalized Method of Moments Specification Testing," *Journal of Econometrics*, September 1985, *29*, 229–56.

Rosen, Sherwin and Taubman, Paul, "Changes in Life Cycle Earnings: What Do Social Security Data Show?" *Journal of Human Resources*, Summer 1982, *17*, 321–38.

Schwartz, Saul, "The Relative Earnings of Vietnam and Korean-Era Veterans," *Industrial and Labor Relations Review*, July 1986, *39*, 564–72.

Tarr, Curtis W., *By the Numbers: The Reform of the Selective Service System 1970–72*, Washington: National Defense University Press, 1981.

Taussig, Michael K., "The Rationale for Veterans Benefits," in *Those Who Served: Report of the 20th Century Fund Task Force on Policies Toward Veterans*, New York: The 20th Century Fund, 1974.

Wald, Abraham, "The Fitting of Straight Lines If Both Variables Are Subject to Error," *Annals of Mathematical Statistics*, September 1940, *11*, 284–300.

Welch, Finis, "The Effect of Cohort Size on Earnings: The Baby Boom Babies Financial Bust," *Journal of Political Economy*, October 1979, Part 2, *87*, S69–S97.

Willis, Robert J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in O. Ashenfelter and R. Layard, eds., *The Handbook of Labor Economics*, Amsterdam: North-Holland, 1986.

Council of Economic Advisors, *Economic Report of the President*, February 1988, Washington: USGPO, 1988.

Selective Service System, Office of Public Affairs, "*A Short History of the Selective Service System*," Washington: USGPO, 1986.

______, *Semiannual Report of the Director of Selective Service for the Period July 1 to December 31, 1970*, Washington: USGPO, 1971.

______ (1969–1973), *Semiannual Report of the Director of Selective Service*, selected issues, Washington: USGPO, 1969–73.

U.S. Department of Health and Human Services, Social Security Administration, *Social Security Bulletin Annual Statistical Supplement*, Washington: USGPO, 1987.

U.S. Department of Commerce, Bureau of the Census, *Survey of Income and Program Participation: Wave I Rectangular File*, Washington: USGPO, 1985.

Veterans Administration (1981a), Reports and Statistics Service, Research Monograph 15: *Educational and Income Characteristics of Veterans, March 1979*, Washington: USGPO, August 1981.

______, (1981b), Office of Reports and Statistics, *Data on Vietnam Era Veterans*, Washington: USGPO, September 1981.

______, (1984), Administrator of Veterans Affairs, *Annual Report 1983*, Washington: USGPO, 1984.