NBER WORKING PAPERS SERIES

ESTIMATING THE PAYOFF TO SCHOOLING USING THE VIETNAM-ERA DRAFT LOTTERY

Joshua D. Angrist

Alan B. Krueger

Working Paper No. 4067

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1992

We are grateful to Kevin McCormick, Ronald Tucker, and Greg Weyland for creating a special Current Population Survey extract for us. We are also grateful to Dean Hyslop for excellent research assistance and to Michael Boozer for helpful comments. The data used in this paper will be made available to other researchers. This research was supported by the National Science Foundation, Grant No. SES-9012149. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #4067
May 1992

ESTIMATING THE PAYOFF TO SCHOOLING USING THE VIETNAM-ERA DRAFT LOTTERY

ABSTRACT

Between 1970 and 1973 priority for military service was randomly assigned to draft-age men in a series of lotteries. Many men who were at risk of being drafted managed to avoid military service by enrolling in school and obtaining an educational deferment. This paper uses the draft lottery as a natural experiment to estimate the return to education and the veteran premium. Estimates are based on special extracts of the Current Population Survey for 1979 and 1981-85. The results suggest that an extra year of schooling acquired in response to the lottery is associated with 6.6 percent higher weekly earnings. This figure is about 10 percent higher than the OLS estimate of the return to education in this sample, which suggests there is omitted-variable bias in conventional estimates of the return to education. Our findings are robust to a variety of assumptions about the effect of veteran status on earnings.

Joshua D. Angrist
Economics Department
Hebrew University
Jerusalem 91905
Israel
and NBER

Alan B. Krueger
Economics Department
Princeton University
Princeton, NJ 08544
and NBER

For many years economists have sought to estimate the monetary return to education without bias from omitted variables that are correlated with educational attainment and with earnings capacity (see Griliches (1977) for a survey of this literature). A positive bias would arise, for example, if individuals with higher earnings capacity obtain more schooling. A variety of econometric techniques have been applied to try to overcome this problem, but the magnitude and importance of omitted-variable bias in the estimated return to education is still an unanswered question. Willis (1986; p. 589), for example, concludes that, "Given the complexity of the issues and the non-representative character of the data sets that have been employed in the literature on ability bias, it is difficult to reach any firm conclusion about the magnitude or even the direction of the bias...."

Concern over potential ability bias in the return to education has been heightened in recent years by evidence that the return to education has increased in the 1980's.[1] Of course, if individuals' schooling levels were randomly assigned, the payoff to education could be consistently measured simply by comparing the earnings of more and less educated individuals. In the absence of random assignment, ability bias can be overcome by using instrumental variables that are correlated with education but have no other effect on earnings. This paper reports our analysis of a natural experiment that generates such instruments. The natural experiment stems from the Vietnam-era draft lottery, which was used to determine priority for induction into the U.S. military between 1970 and 1973.

As a consequence of the draft lottery, the risk of induction into the armed services was randomly assigned on the basis of individuals' birthdays.

---

[1] See, for example, Blackburn and Neumark (1991). For convenience, we follow the practice of the previous literature and label omitted variables in wage equations that are correlated with education as ability. We note that other omitted factors, such as family wealth, might influence educational achievement and earnings.

In certain years, however, men could be deferred or exempted from military service by remaining in school and obtaining an educational deferment. Thus, an individual's draft lottery number affected both the likelihood he would serve in the military and the incentive for him to seek additional schooling. Because past evidence has established a link between military service and civilian earnings (for examples see Berger and Hirsch, 1983, Angrist, 1990, and Angrist and Krueger, 1989), the impact of military service on earnings must be adequately accounted for if the influence of the draft lottery is to be used to identify the effect of schooling on earnings. We develop and implement an estimation strategy that uses draft lottery number dummies as instruments for both the return to education and for the effect of veteran status on earnings.

Our estimation strategy is feasible because young men seeking to avoid the draft had to take draft-avoidance measures (e.g., school attendance) before the highest lottery number to be called for the draft was actually known. Therefore, the lottery has an influence on years of schooling that differs from its influence on veteran status. For example, after conditioning on veteran status, we find that men at high risk of being drafted were 5 percentage points more likely to have attended college. Furthermore, college deferments were curtailed in 1971, changing the relationship between lottery numbers and education over time. Thus, in principle it is possible to use functions of draft lottery numbers as instruments for both veteran status and education.

We analyze data from six March Current Population Surveys (CPS) that were specially prepared for this project by the Census Bureau. These data sets include information on the lottery numbers assigned to men who participated in the Vietnam-era draft lottery. The CPS extracts also contain labor market information for the years 1979 and 1981-1985. Perhaps surprisingly, we find no evidence that conventional ordinary least squares (OLS) estimates of the rate of return to education are biased upwards. In fact, the lottery-based estimate of the rate of return to schooling is slightly higher than the OLS estimate. This finding is robust to a variety of assumptions about the impact of veteran status on earnings. For example, our qualitative conclusions are unchanged if we make a wide range of assumptions about the effect of veteran status on earnings. These results provide support for the view that ability bias is not a significant problem in estimates of the return to education, at least for this cohort of men.

The paper is organized as follows. Section I describes institutional features of the Vietnam-era draft lottery, with particular emphasis on the ways in which the lottery affected school enrollment. Section II describes the data that we analyze. Section III presents our econometric model, and section IV presents our empirical findings. Section V explores the robustness of these findings.

I. Educational Deferments and the Draft Lottery

Each draft lottery held 1970-1973 consisted of the assignment of Random Sequence Numbers (RSN's or "draft lottery numbers") to all dates of birth in the cohort to be called for service.[2] As the year progressed, men were called for induction by RSN, with the lowest numbers called first. At some date during the year, an official RSN ceiling was determined according to Defense Department manpower needs; only men with lottery numbers below the ceiling, referred to here as draft-eligible men, could have been drafted. An essential feature of this process was that for most of the year during which draft registrants were at risk of being drafted, they did not know what the

[2]The 1970 lottery, which included men born between 1944 and 1950, did not generate a truly random sequence of lottery numbers. For example, Fienberg (1971) shows that the mean lottery number increased with month of birth. This is unlikely to seriously affect our results because month of birth is probably uncorrelated with the residual in the earnings equation.

ultimate lottery number ceiling would be. Therefore, even for men with relatively high numbers, there was a strong incentive for those who wanted to avoid military service to take "preventive action" by obtaining a draft deferment. Men with low lottery numbers were also substantially more likely to voluntarily enlist, thereby improving their conditions of service (see Angrist, 1991).

Separate birth cohorts were at risk of being drafted in each lottery. The 1970 lottery involved men born in the years 1944-1950, the 1971 lottery involved men born in 1951, the 1972 lottery involved men born in 1952 and so on, through 1975. Few men were actually drafted after 1972, however. And in July of 1973 congressional authority to induct servicemen expired and was never renewed.

A few months before the year in which a cohort was slated for induction, lottery numbers were assigned to dates of birth in a televised national drawing. Final selection from the pool of draft-eligible men was largely based on the screening process employed by local boards and by the elimination of men with deferments (Tarr 1981). According to the "channeling" policy, individuals were granted deferments for a variety of educational, occupational, or family reasons. Deferred men retained the liability implied by the RSN attached to their date of birth upon expiration of their deferment. For example, a deferred college student with a low number who was born in 1950 would face a high risk of being drafted if he left school in 1972. High school students were also automatically deferred as long as they remained in school.

A chronology of key events involving Selective Service procedures is presented in Table 1. The chronology illustrates the uncertainty young men faced after they were assigned a random sequence number. Although official induction ceilings were often reached midway through the year, the highest RSN to be called was generally not known until later in the year. For example, the 1971 ceiling for induction (125) was first called in May. But 125 was not officially declared to be the induction ceiling until October, and men with numbers as high as 170 were called for pre-induction physicals. Similarly, although no one was inducted after 1972, men born in 1953 with lottery numbers below 100 were eligible (but not necessarily called) for administrative processing in 1973.

Because of uncertainty regarding the risk of induction associated with particular lottery numbers, the pattern of variation in deferment behavior with lottery numbers is likely to differ from the pattern of variation in the probability of veteran status. For example, consider the position of a man born in 1951 with a lottery number between 125 and 195. This man was first at risk of being drafted in 1971. In the preceding year, RSN 195 was the highest number called. The draft eligibility ceiling for the 1971 lottery was set at 125, but this was not announced until October. Thus, a student in this situation would have an incentive to stay in school to avoid military service, even though it turned out that he could not have been drafted given his lottery number. Circumstances like these lead to a relationship between schooling and lottery numbers that differs from the relationship between the probability of serving in the military and lottery numbers.

Table 1 also reports the number of men with student deferments at various dates in the lottery period. Although new graduate student deferments were eliminated in 1967 and new undergraduate deferments were eliminated in December 1971, there were still over one million deferred students at the beginning of 1972. This is because students who were already deferred preceding the rule changes were generally allowed to keep their deferments (Baskir and Strauss 1978). In 1970 and 1971 -- the peak years for inductions during the lottery -- there were over 2 million deferred high school and college students. Among men found physically and mentally acceptable for service, the vast majority of deferments were obtained for educational reasons.[3]

Time Series Evidence

For most of the post World War II period, school enrollment rates of 18 and 19 year olds have been rising. As shown in Figure 1, however, the school enrollment rate increased rapidly in the mid to late 1960's, with male enrollment peaking in 1968. The 1968 enrollment peak and subsequent steep decline in the early 1970's is often attributed to educational deferments during the Vietnam-era draft (e.g., Fuchs 1983, p. 93). Baskir and Strauss (1978) argue that during the Vietnam era, college enrollment rates were 6-7% higher than normal because of the draft, and Singer (1989) argues that the draft accounts for exceptionally strong growth in the number of applicants to medical school during the 1968-74 period. Finally, Bowen and Rudenstine (1992) present evidence that the elimination of graduate school deferments contributed to the decline in Ph.D.'s awarded in the 1970s.

In spite of considerable circumstantial evidence regarding the relationship between the draft and higher education, it is difficult to prove that the draft accounts for the 1960's enrollment spike. For example, the observation that enrollment rates of women follow a pattern similar to those of men is prima facie evidence that the change in enrollment rates was not entirely due to draft avoidance. Moreover, Freeman (1975) presents evidence that the bulk of the 1960's increase in education is due to economic factors. Nevertheless, we find evidence that the draft did affect school enrollment

---

[3] 162,746 men were drafted in 1970, 94,092 men in 1971, 49,514 men in 1972 and only 646 in 1973. 382,010 men were drafted in 1966, at the peak of Vietnam Era accessions (Selective Service System 1986). Of 26.8 million men of draft age during the Vietnam Era (1964-75), 8.7 million voluntarily enlisted and 2.2 million were drafted (Baskir and Strauss 1978).

decisions for many draft-eligible individuals. In particular, we use micro data to document a small but statistically significant relationship between draft lottery numbers and educational attainment.

II. Data

Our analysis requires a data set containing information on lottery numbers, education, veteran status, and earnings. Draft lottery numbers can be determined on the basis of date of birth from published tables (Selective Service System 1969-73). However, date of birth is usually excluded from public-use micro data sets such as the Current Population Survey (CPS) and Population Census because of concern that it compromises respondent anonymity. Consequently, this information is not available in most of the data sets analyzed by labor economists.[4]

Although not available on CPS public-use tapes, date of birth is recorded on the "Control Cards" used by the Census Bureau for record keeping purposes. The Census Bureau cannot release data containing individuals' exact date of birth, but we have worked with the Bureau to create six special Current Population Survey files that include a new variable constructed by recoding men's birthdays into 14 groups of lottery numbers. These special extracts contain a lottery number recode variable that indicates 13 groups of 25 consecutive lottery numbers (RSN 1-324), and a final group of 40 consecutive numbers (RSN greater than 325). Because several birthdays are grouped into lottery number ranges, respondent anonymity is not jeopardized in these files. A variable indicating year of birth was also added to the tape. The final data set includes all men born 1944-53 in rotation

[4]One exception is the National Longitudinal Survey, which contains information on date of birth. But this data set is too small to yield precise estimates for our purposes.

groups 5-8 of the March 1979 CPS, and in rotation groups 5-8 of the March 1981-85 CPS's. There are 30,967 men in this sample. After we eliminate men who have missing values or who did not work during the year, the sample size is 26,119.

III. Lottery-Based Estimation

Our interest is in estimating parameters of the following equation:

$$(1) \qquad y_i = s_i\rho + v_i\gamma + X_i\beta + \epsilon_i$$

where y_i is i's log earnings, s_i is years of schooling, v_i is veteran status, and X_i is a set of covariates such as region of residence. ϵ_i is a regression error with scalar covariance, that may be correlated with both s_i and v_i. The parameter ρ is commonly referred to as the return to education, and γ the veteran premium. For purposes of exposition we ignore the covariates, X_i, but we include them in the equations we estimate.

Ordinary Least Squares (OLS) estimates of the return to education, ρ, may be biased because of correlation between s_i and components of ϵ_i, such as ability. The draft lottery, however, provides instrumental variables that are correlated with education and with veteran status but likely to be uncorrelated with other variables related to earnings.

As in all Two-Stage Least Squares (2SLS) problems, precision of the estimates in the second stage depends in part on the correlation between instruments and endogenous regressors in the first stage equation. We begin analysis of the effect of lottery numbers on education by assuming that veteran status is orthogonal to the regression error in (1). The question of whether lottery numbers may be used as an instrument for education then hinges on the correlation between education and functions of lottery numbers, conditional on veteran status. We work with 13 dummies indicating 25 consecutive lottery numbers because this is the least restrictive specification for the data we have.[5] The 13 dummies are interacted with year-of-birth dummies to allow the relationship between lottery numbers and schooling to vary by cohort. In the analysis where both veteran status and education are treated as endogenous regressors, the use of unrestricted dummies provides enough instruments to identify the effects of both variables.

Let z_i be the row vector of lottery number dummies that is used as an instrument for individual i. Ignoring other covariates, the following is our first stage equation for schooling, assuming veteran status is exogenous:

$$s_i = v_i\delta + z_i\alpha + \eta_i, \tag{2}$$

where δ is a coefficient that measures the effect of veteran status on education.[6]

The first three columns of Table 2 report OLS estimates of equation (2). The dependent variable is years of completed education. To illustrate the impact of lottery numbers on educational attainment for each birth year, the table reports coefficients and standard errors for three broad lottery group dummies: one for RSN 1-75; one for RSN 76-150; and one for RSN 151-225. The omitted lottery group is RSN 226-365, which is the least likely group to be drafted. Each RSN dummy is interacted with year-of-birth dummies. Results are presented separately for each year of birth, and a test of the joint significance of the lottery number dummies is presented in the last row of the table. The estimated equation also includes a veteran status dummy, two race dummies, a central city dummy, a balance of SMSA dummy, a marital status

---

[5]The omitted group contains numbers above 325.

[6]For example, veteran status may be related to education because of tuition assistance from the GI Bill.

dummy, 5 year dummies, 9 year-of- birth dummies and 8 region dummies.[7]

The estimates suggest that the lottery groups have a significant relationship with educational attainment for certain birth years. For example, men in the 1946 cohort who were assigned low lottery numbers obtained about .3 years more education than men who were assigned to the highest lottery group. Overall, an F-test of the hypothesis that the lottery groups are unrelated to educational attainment has a p-value of 7 percent. It is difficult to discern a clear or consistent pattern in the lottery group coefficients, however. Low lottery numbers are associated with less schooling for the 1944 cohort, but with more schooling for the 1945 and 1946 cohorts. And the lottery numbers appear to have little or no effect on education for cohorts born after 1947. The changing pattern of the relationship between lottery numbers and education may reflect changes in deferment policy over time and changes in individuals' perceptions about the risk of being drafted over time.

We have also estimated equation (2) using other measures of educational attainment as the dependent variable. For example, linear probability models for whether individuals attended some college indicate that men in the lowest lottery number group are 5 percentage points more likely to attend college than men in the highest lottery number group.

The estimates for education in Table 2 are of interest if veteran status is an exogenous regressor in equation (1). However, Angrist (1989) reports evidence suggesting that veteran status may be correlated with the regression error in (1), and should therefore be treated as an endogenous variable. Draft lottery numbers provide natural instruments for veteran status as well as for years of education. Consistent estimation of the returns to education therefore turns on the ability to use functions of lottery numbers as instruments for both veteran status and education.

---

[7]These are the same covariates used in the 2SLS models below.

The problem of identification with two endogenous regressors and one set of lottery dummies as instruments can be analyzed as follows. Re-write equation (2) as:

$$s_i = v_i\pi + \eta_i, \tag{3}$$

where $\eta_i = z_i\alpha + \nu_i$. Let $P_z = Z(Z'Z)^{-1}Z'$ be the projection matrix associated with the instrument matrix, Z. If both s_i and v_i are endogenous, 2SLS estimates of (1) are the same as OLS estimates of

$$Y = P_zS\rho + P_zV\gamma + \epsilon, \tag{4}$$

where capital letters denote vectors. Therefore, the TSLS estimate of ρ is equivalent to the coefficient estimate from a bivariate OLS regression of Y on

$$[I - P_zV(V'P_zV)^{-1}V'P_z]P_zS = P_z[I - V(V'P_zV)^{-1}V'P_z]S = P_zQ_{zv}S,$$

(where $Q_{zv} = [I - V(V'P_zV)^{-1}V'P_z]$). $Q_{zv}S$ is the residual after estimating equation (3) by 2SLS using z_i as instruments for v_i.

Thus, if V and S are both endogenous regressors, ρ is identified only when Z is be correlated with $Q_{zv}S$. That is, after subtracting off fitted values of veteran status -- where the fitted values are from a regression of veteran status on Z -- schooling must still be correlated with Z. To illustrate this, we first estimate the relationship between schooling and veteran status and other covariates by 2SLS, using lottery dummies (and covariates) as instruments for veteran status. We then calculate the residuals from this equation, and regress these residuals on the lottery dummies (Z) and the covariates.

The estimates are presented in columns (4)-(6) of Table 2. Results from regressions of the 2SLS residuals from (3) on lottery dummies are quite similar to the results that simply control for veteran status. The prob-value for an F-test of the joint significance of the lottery dummies is 8 percent, which is only slightly higher than for the specification that treats veteran status as exogenous. The pattern of coefficients on lottery group dummies is also similar.

IV. Empirical Results

Table 3 reports our main set of estimates of equation (1). The sample used to estimate these models consists of men born between 1944 and 1953.[8] Column (1) reports results for a model in which years of education is treated as an endogenous regressor and veteran status as an exogenous regressor. The equation is estimated by 2SLS, using 130 lottery-by-year-of-birth dummies as (excluded) instrumental variables. In column (2) veteran status is modelled as an endogenous regressor and education as an exogenous regressor. Again, the model is identified by the exclusion of dummy variables for lottery number groups. The model in column (3) treats both education and veteran status as endogenous regressors, and instruments for both variables with the lottery number dummies. Finally, the model in column (4) is estimated by OLS. All of the equations also include several covariates that are reported in the table.

We first discuss the estimates of the rate of return to education. The estimated return to education is 6.6 percent (t-ratio=4.4) if education is the only endogenous regressor. If education and veteran status are both treated as endogenous regressors, the rate of return to education is still 6.6 percent. In this sample the OLS estimate of the return to education is 5.9 percent, slightly less than the 2SLS estimate.

---

[8]Earnings were converted to 1978 dollars using the CPI. The weekly wage is the ratio of annual earnings in the preceding calendar year to weeks worked in the same year. We eliminate individuals whose weekly wage is greater than or equal to $2,500 or is less than $25.

The 2SLS and OLS estimates are not significantly different from each other. Nevertheless, it is interesting to compare the magnitude of the 2SLS to the OLS estimate of the return to education. Suppose the only specification error in the estimated earnings function is mismeasurement of education. If the measurement error in education is uncorrelated with true education and uncorrelated with the other regressors, then the ratio of the OLS to the 2SLS estimates of the coefficient for education has probability limit equal to the ratio of the variance in true education to the variance in reported education.[9] This quantity is often called a reliability ratio. In this case, the reliability ratio is about .89, which is close to estimates based on re-survey data. For example, Siegel and Hodge (1968) estimate that the reliability ratio for education is .933, based on data from the 1960 Census and the 1960 Census post-enumeration survey. And Bielby, Hauser, and Featherman (1977) report reliability ratios of .801, .838, and .921 using data from the March 1973 CPS, Occupational Changes in a Generation survey, and a re-survey sample.[10] The unweighted mean of these four estimates is .873.

<u>Veteran Premium</u>

The estimated veteran premium is statistically insignificant in the 2SLS models reported in Columns 2 and 3. The coefficient is slightly positive, but the standard error is relatively large. In the OLS model, veteran status also has a small, positive coefficient. Apparently, the reason why the education variable is robust to the inclusion of veteran status is because veteran

---

[9]This statement assumes that true schooling is orthogonal to the other regressors.

[10]These reliability ratios are the correlations between education for individuals who were sampled at two different dates. Assuming that measurement error is i.i.d. and uncorrelated with true education, these correlations identify the reliability ratio.

status has an insignificant effect on earnings in this data set. The small positive estimate of the veteran premium is a contrast with the past literature, which typically finds that Vietnam-era veterans earn less than nonveterans.[11] A possible explanation for these contrasting findings is that we have analyzed data for a wide age group to increase our sample. For example, Angrist (1990) estimates the Vietnam-era premium for men born 1950-1953, and focuses on annual earnings instead of weekly earnings.

To further explore the effect of veteran status on earnings, Table 4 reports 2SLS and OLS estimates of wage equations using the log annual wage as the dependent variable, and presents additional estimates for the subsample of men born between 1950 and 1953. Panel A reports estimates for the full sample (men born between 1944 and 1953) using the log of annual earnings rather than weekly earnings as the dependent variable.[12] As to be expected, the rate of return to education is larger in models that use annual earnings than in models that use weekly earnings. The veteran premium, however, is negative in the models that use annual earnings and instrument for veteran status. Panel B restricts the sample to men born between 1950 and 1953, and uses annual earnings as the dependent variable. Models in Panel C also use the narrower sample, but use the log of the weekly wage as the dependent variable. Estimates for the 1950-53 birth cohorts using the log of annual earnings as the dependent variable show a negative veterans premium which

---

[11]For examples of this literature, see Angrist (1989), De Tray (1982), Rosen and Taubman (1982), Crane and Wise (1987). Using CPS data, however, Berger and Hirsch (1983) find an insignificant earnings premium for Vietnam-era veterans compared to similar aged nonveterans.

[12]The sample size is larger in Panel A of Table 4 than in Table 3 because we do not eliminate observations that have outliers or missing values for their weekly wage when we use annual earnings as the dependent variable. The results are not very different if we restrict the sample to the same observations used to estimate Table 3.

is close in magnitude to Angrist's (1990) estimate based on grouped Social Security data.

It is also worth noting that the 2SLS estimate of the return to education remains substantial if the sample is restricted to men born 1950-53. For example, the estimate of the return to schooling of .075 in column 3 of Table 4 is close to the corresponding estimate of .066 for the full sample in column 3 of Table 3. But the standard errors are about twice as large for the 1950-53 subsample. Nevertheless, the 1950-53 cohort is of particular interest because these men were close to college-entry age when they participated in the draft lottery in the early 1970's.

V. Robustness of the Estimated Rate of Return to Schooling

One potential problem with using draft lottery numbers as instruments for education and veteran status is that lottery numbers may also be related to other forms of draft avoidance. For example, some men who were assigned low lottery numbers fled to Canada or intentionally injured themselves to avoid military service. The possibility that such behavior also affects earnings cannot be ruled out. But we suspect that such behavior is unlikely to meaningfully bias our estimates. Most deferments in the Vietnam era were due to failure to meet physical and mental standards. These standards were high enough to eliminate the majority of draft-eligible young men without further action on their part (Tarr, 1981). For example, wearing moderately strong corrective lenses was often grounds for deferment. Among those not eliminated in this manner, education was the major source of deferment in the 1970's.

The last row of Table 3 reports the p-value for a test of instrument-error orthogonality. Specification error caused by other potential forms of draft avoidance might be picked up by these over-identification tests. In all cases, however, the prob-values are far greater than conventional significance levels.

Identification of Education and Veteran Coefficients

The success of our identification strategy depends on whether the relationship between lottery numbers and veteran status is linearly independent of the relationship between lottery numbers and education. We consider this to be mainly an empirical issue that is resolved by the standard errors of the estimates, which in this case suggests there is not an identification problem. Nevertheless, the robustness of our estimate of the rate of return to education can be probed further by making alternative assumptions about the magnitude of the coefficient on veteran status in equation (1). If we constrain the veteran coefficient (γ) to equal a specific value, the education coefficient is identified even if education and veteran status are identical functions of lottery numbers. Our approach is to vary γ over a wide range of values to explore the robustness of the 2SLS estimate of ρ.

Consider the equation:

(5) $$y_i = s_i\rho + \{v_i\gamma + \epsilon_i\},$$

where $v_i\gamma$ is now an omitted variable in the compound error term. Then a straightforward extension of the omitted-variable-bias formula gives the 2SLS estimate of ρ as a function of γ:

(6) $$\hat{\rho}(\gamma) = \hat{\rho}_0 - \hat{\theta}\,\gamma,$$

where $\hat{\rho}_0$ is the 2SLS estimate of $\hat{\rho}$ imposing $\gamma = 0$, and $\hat{\theta}$ is the 2SLS estimate of θ in an auxiliary regression of veteran status on schooling:[13] $v_i = s_i\theta + \nu_i$. The auxiliary 2SLS equation uses lottery dummies as instruments to identify the parameter θ.

To implement this idea, we parametrically vary the coefficient on veteran status between -.30 and +.30. This is a wide range, and it is extremely likely that the true effect of veteran status on civilian earnings for this cohort falls within this range.

Figure 2 presents a plot of 2SLS estimates of the rate of return to education under alternative assumptions about the magnitude of the veteran coefficient. The figure shows that the return to education is remarkably robust to a range of assumptions about the magnitude of the veteran premium. For example, if the veteran premium is constrained to equal -.30 the 2SLS estimate of the return to education is .064 with a standard error of .016, and if the veteran premium is constrained to equal +.30, the 2SLS estimate of the return to education is .068 with a standard error of .016. In both cases the 2SLS estimate exceeds the OLS estimate of the return to education.[14] The return to schooling is robust to extreme assumptions about the effect of veteran status because $\hat{\theta}$ is small, only -.0066.

---

[13]Proof of (6): Alternative estimates of $\hat{\rho}(\gamma)$ may be computed by 2SLS estimation of $\hat{\rho}$ after subtracting $v_i\gamma$ from the dependent variable:

$$(y_i - v_i\gamma) = s_i\rho + \epsilon_i.$$

Therefore,

$$\hat{\rho}(\gamma) = (S'P_zS)^{-1}S'P_zY - [(S'P_zS)^{-1}S'P_zV]\gamma = \hat{\rho}_0 - \hat{\theta}\gamma.$$

Notice that if fitted values of v_i and s_i (from regressions on z_i) are orthogonal, θ is zero and the 2SLS estimate of ρ is asymptotically unbiased.

[14]The OLS estimate of the return to education only varies between .058 and .059 as the veteran premium is varied between -.30 and +.30.

V. Conclusion

This paper examines the civilian earnings experience and education of men who were subject to the Vietnam-era draft lottery. Because men who were concerned about the risk of induction could reduce their chances of being drafted by seeking a student deferment, the draft lottery generated exogenous variation in years of schooling. We exploit this variation to estimate the payoff to additional years of schooling. In addition, we use the natural experiment of the draft lottery to estimate the effect of veteran status on civilian earnings.

Our main finding is that the variation in schooling generated by the draft lottery causes about a 6.5 percent increase in weekly earnings per year of additional schooling. This is somewhat higher than the conventional OLS estimate of 5.9 percent for this sample. Furthermore, the lottery-based estimate of the return to schooling is robust to alternative assumptions about the magnitude of the veteran premium.

We conclude from these results that there is little evidence of positive ability bias in conventional estimates of the return to education. In fact, our lottery-based estimates of the payoff to schooling suggest that conventional OLS estimates may be somewhat understated. Moreover, estimates of the extent of measurement error in education based on re-interview data suggest that the downward bias in the OLS estimate of the return to education is of the same order of magnitude that would be expected if measurement error in education were the only source of specification error in the human capital earnings function. These findings parallel Angrist and Krueger's (1992) estimates of the return to schooling based on compulsory schooling laws, and Ashenfelter and Krueger's (1992) estimates of the return to schooling using twin methods. Similarly, Griliches (1977; p. 18) concludes that the bias in typical OLS estimates of the return to education is "either nil or negative." It is our conclusion that the consistent finding of no omitted-variable bias in OLS estimates of the return to education should be a feature of theoretical models of education.[15]

[15]Weiss (1971) shows that, depending on the relationship between ability and earning capacity, there may be any relationship between "ability" and the optimal level of schooling in the human capital model.

REFERENCES

Angrist, J.D., "Using the Draft Lottery to Measure the Effects of Military Service on Civilian Earnings," Research in Labor Economics, Vol. X, edited by Ronald Ehrenberg (Greenwich, CT: JAI Press, Inc., 1989).

Angrist, J.D., "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," American Economic Review, 80, No. 3, June 1990, pp. 313-336.

Angrist, J.D., "The Draft Lottery and Voluntary Enlistment in the Vietnam Era," Journal of the American Statistical Association 86, September 1991, pp. 584-595.

Angrist, J.D. and Alan Krueger, "Why Do World War Two Veterans Earn More than Nonveterans?", NBER Working Paper No. 2991, May 1989.

Angrist, J.D. and Alan Krueger, "Does Compulsory Schooling Affect Schooling and Earnings?" Quarterly Journal of Economics 106, November 1991, pp. 979-1014.

Ashenfelter, Orley, and Alan Krueger, "Estimates of the Economic Return to Schooling from a New Sample of Twins," mimeo., Princeton University, 1992.

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

Berger, M.C. and B.T. Hirsch, "The Civilian Earnings Experience of Vietnam-Era Veterans," The Journal of Human Resources 18, No. 4, Fall 1983, pp. 455-479.

Bielby, William, Robert Hauser, and David Featherman, "Response Errors of Non-Black Males in Models of the Stratification Process," in D.J. Aigner and A.S. Goldberger, editors, Latent Variables in Socio-conomic Models (Amsterdam: North Holland, 1977), pp. 227-251.

Blackburn, Mckinley, and David Neumark, "Omitted-Ability Bias and the Increase in the Return to Schooling," NBER W.P. No. 3693, May 1991.

Bowen, William G., and Neil L. Rudenstine, In Pursuit of the PHD (Princeton, NJ: Princeton University Press, 1992).

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

De Tray, Dennis, "Veteran Status as a Screening Device," American Economic Review 72, March 1982, pp. 133-42.

Feinberg, Stephen, "Randomization and Social Affairs: The 1970 Draft Lottery," Science 171, January 1971, pp. 255-308.

Freeman, Richard B., "Overinvestment in College Training?," Journal of Human Resources 10, 1975, pp. 287-311.

Fuchs, Victor, How We Live: An Economic Perspective on Americans from Birth to Death, Cambridge: Harvard University Press, 1983.

Griliches, Zvi, "Estimating the Returns to Schooling - Some Econometric Problems," Econometrica 45 (January 1977): 1-22.

Hungerford, Thomas and Gary Solon, "Sheepskin Effects in the Returns to Education," Review of Economics and Statistics 66, February 1987, pp. 175-77.

Rosen, Sherwin, and Paul Taubman, "Changes in Life Cycle Earnings: What Do Social Security Data Show?" Journal of Human Resources 17, no. 3, Summer 1982, pp. 321-38.

Selective Service System, "Semiannual Report of the Director of Selective Service" (Selected issues, 1969-1973), Washington DC: U.S. Government Printing Office.

Selective Service System, "A Short History of the Selective Service System," Office of Public Affairs, Selective Service System National Headquarters, Washington, DC: U.S. Government Printing Office, 1986.

Siegel, Paul, and Robert Hodge, "A Causal Approach to the Study of Measurement Error," in Hubert Blalock and Ann Blalock, editors, Methodology in Social Research (New York: McGraw Hill Book Co., 1968), pp. 28-59.

Singer, Allen, "The Effect of the Vietnam War on Numbers of Medical School Applicants," Academic Medicine 64, 1989, pp. 567-73.

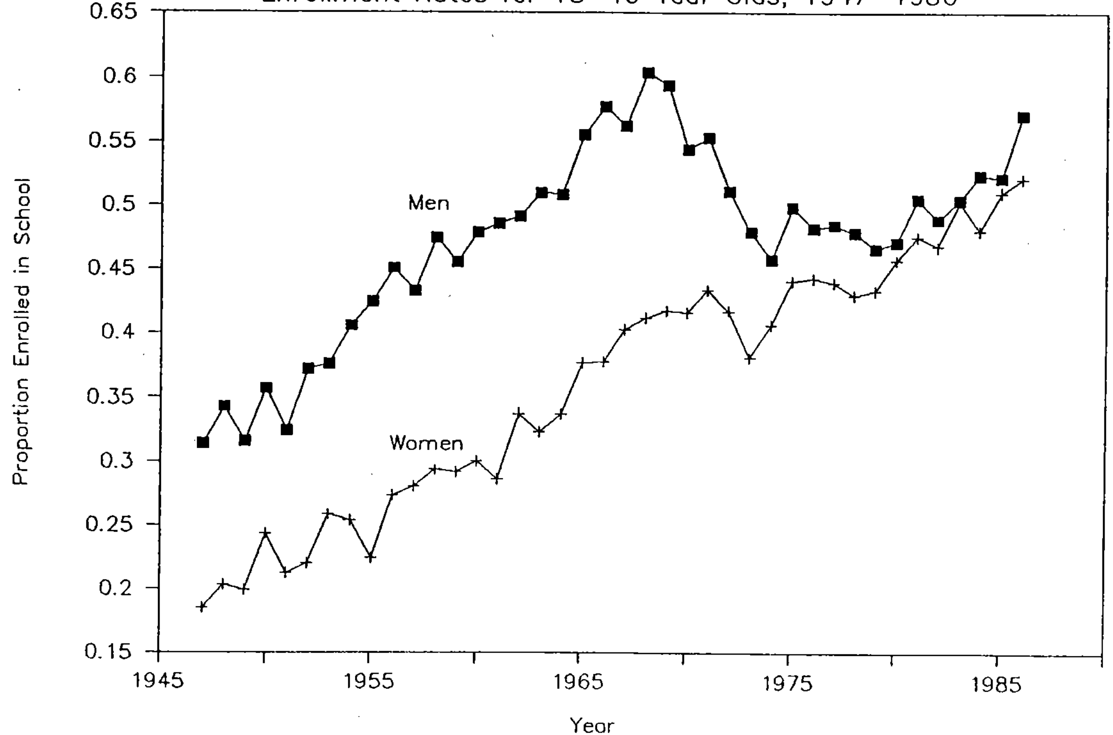
Tarr, C.W., By the Numbers: The Reform of the Selective Service System 1970-72, Washington DC: National Defense University Press, 1981.

Weiss, Yoram, "Ability and the Investment in Schooling: A Theoretical Note on J. Mincer's 'Distribution of Labor Income'," Journal of Economic Literature 9, June 1971, pp. 459-461.

Willis, Robert J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," Chapter 10 in the Handbook of Labor Economics, Volume I, Edited by O. Ashenfelter and R. Layard, Amsterdam: Elsevier Science Publishers BV, 1986.

Figure 1

Enrollment Rates for 18–19 Year Olds, 1947–1986



Sources: Labor Force Statistics Derived from the CPS, 1947–1986; School Enrollment, P–20 Series, No. 409 and 429.

Figure 2

Estimated Rate of Return to Education Varying Constraints on Veteran Status

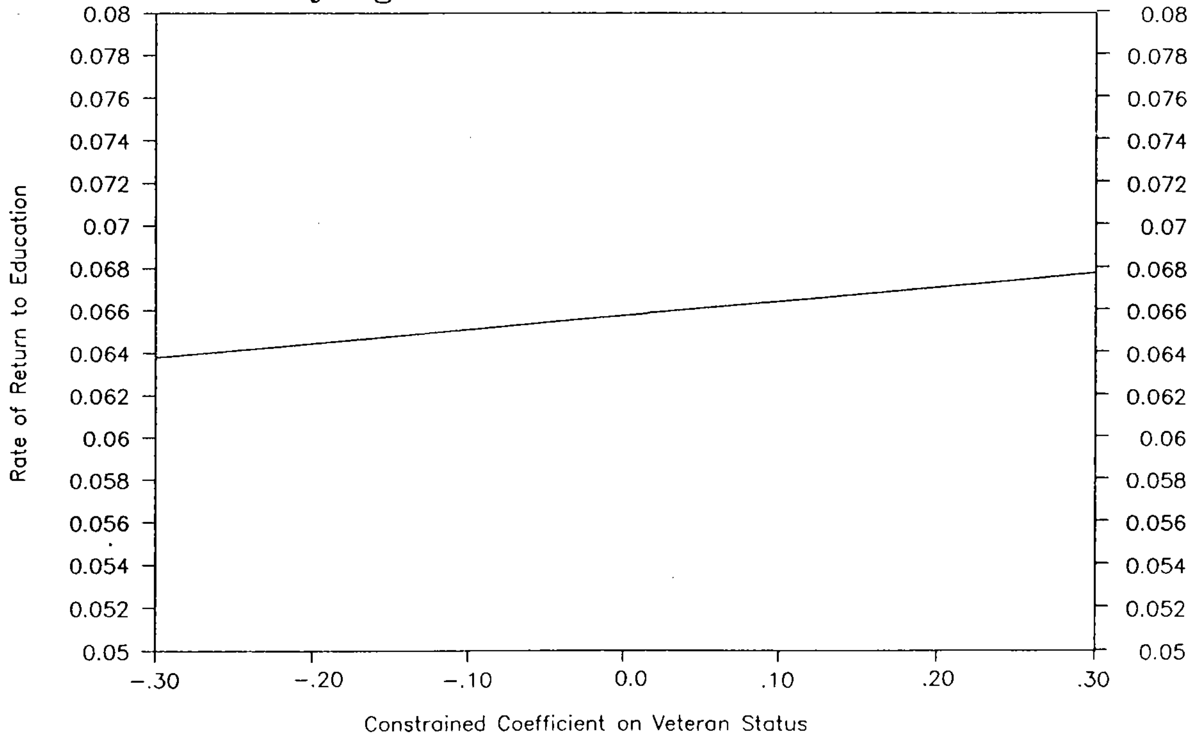


Table 1. Random Selection and Deferment 1967-73

Lottery	Event and Students Deferred	Date
1967	1967 Selective Service Act	June 30, 1967
	Most new graduate deferments ended. Currently deferred allowed to end term or semester. Grad students teaching full time granted occupational deferments. Undergraduate and high school deferments continue.	
	number deferred: 1.703 million college .526 million high school	June 30, 1967
1970	Random Selection for Birth Cohorts 1944-50	December 1, 1969
	number deferred: 1.843 million college .418 million high school	January 1, 1970
	RSN 1-30 Called	January 1970
	RSN 1-60 Called	February
	RSN 1-90 Called	March
	RSN 1-115 Called	April
	RSN 1-145 Called	May
	RSN 1-170 Called	June
	RSN 1-190 Called	July
	RSN 1-195 Called	August - December
	CEILING: 195	
1971	Random Selection for Birth Cohort 1951	July 1, 1970
	number deferred: 1.699 million college .562 million high school	June 30, 1970
	RSN 1-100 Called	January - April 1971
	RSN 1-125 Called	May - December
	CEILING ANNOUNCED: 125	October 1970
	1971 Draft Reform Amendments	September 28, 1971

- Continued -

Table 1 — Continued

Lottery Event and Students Deferred		Date
	No new undergraduate deferments. Previously deferred allowed to continue in school. High school deferments continue.	December 10, 1971
1972	Random Selection for Birth Cohort 1952	August 5, 1971
	number deferred: .881 million college .232 million high school	December 31, 1971
	RSN 1-15 called RSN 1-75 called RSN 1-95 called	April - June 1972 September October - December
	CEILING ANNOUNCED: 95	September
1973	Random Selection for Birth Cohort 1953	February 2, 1972
	number deferred: .454 million college .030 million high school	June 30, 1972
	Secretary of Defense announces no further draft calls foreseen	January 27, 1973
	Processing, examination and induction suspended	February 14, 1973
	Processing of suspended inductions resumes	April 4, 1973
	Orders issued to terminate all inductions on July 1, 1973	May 15, 1973
	Induction authority expires (for those never deferred)	July 1, 1973

Notes: In 1973, registrants with numbers below 100 were eligible for administrative processing. In 1974, the ceiling for administrative processing was 95. The last lottery was held in 1975, for men born in 1956. Sources: Selective Service System (1969-73, 1986), Baskir and Strauss (1978).

Table 2: Lottery Numbers and Educational Attainment

Year of Birth	Veteran Status Exogenous[a,c] RSN 1-75 (1)	76-150 (2)	151-225 (3)	Veteran Status Endogenous[b,c] RSN 1-75 (4)	76-150 (5)	151-225 (6)
1944	-.194	-.563	-.471	-.197	-.556	-.485
	(.171)	(.174)	(.177)	(.171)	(.174)	(.177)
1945	.375	.230	.457	.378	.235	.460
	(.175)	(.170)	(.175)	(.175)	(.170)	(.175)
1946	.301	.332	.298	.288	.328	.303
	(.157)	(.156)	(.163)	(.157)	(.156)	(.163)
1947	.003	-.228	.008	-.001	-.232	-.012
	(.151)	(.151)	(.147)	(.151)	(.151)	(.148)
1948	-.003	.068	-.004	-.012	.050	-.018
	(.156)	(.152)	(.153)	(.156)	(.152)	(.153)
1949	.057	.091	.079	.031	.076	.082
	(.155)	(.153)	(.150)	(.155)	(.153)	(.150)
1950	-.078	.097	-.131	-.134	.056	-.158
	(.151)	(.152)	(.152)	(.151)	(.152)	(.153)
1951	.358	.160	.139	.312	.122	.133
	(.147)	(.151)	(.146)	(.150)	(.151)	(.146)
1952	-.023	.025	.080	-.084	.011	.081
	(.147)	(.149)	(.149)	(.150)	(.149)	(.149)
1953	-.026	.104	.151	-.035	.102	.149
	(.149)	(.148)	(.150)	(.149)	(.148)	(.150)
P-value for joint F-test		.071			.080	

Notes:

a. Dependent variable is years of schooling.

b. Dependent variable is years of schooling after removing the effect of predicted veteran status and covariates.

c. Covariates are: 2 race dummies, central city dummy,balance of SMSA dummy, marriage dummy, 5 year dummies, 9 year-of-birth dummies, and 8 region dummies. Veteran status is also a covariate when it is treated as exogenous. Sample size is 25,781. Standard errors are shown below the coefficients.

Table 3

2SLS and OLS Wage Equation Estimates
Dependent Variable: Log Real Weekly Wage
Men Born 1944 - 1953

Independent Variables	Mean [SD]	2SLS (1)	2SLS (2)	2SLS (3)	OLS (4)
Intercept	—	4.551 (0.199)	4.639 (0.033)	4.543 (0.201)	4.647 (0.025)
Education	13.39 [2.92]	0.066 (0.015)	0.059 (0.001)	0.066 (0.015)	0.059 (0.001)
Veteran Status	0.34 [0.43]	0.022 (0.007)	0.041 (0.055)	0.042 (0.055)	0.022 (0.007)
Black	0.07 [0.26]	-0.161 (0.017)	-0.167 (0.013)	-0.162 (0.017)	-0.167 (0.013)
Hispanic and Other Races	0.03 [0.18]	-0.121 (0.021)	-0.113 (0.020)	-0.118 (0.023)	-0.116 (0.019)
Central City	0.23 [0.42]	0.016 (0.010)	0.019 (0.009)	0.017 (0.011)	0.018 (0.009)
Balance of SMSA	0.32 [0.47]	0.131 (0.013)	0.136 (0.008)	0.131 (0.013)	0.136 (0.008)
Married with Spouse Present	0.75 [0.43]	0.194 (0.008)	0.192 (0.008)	0.193 (0.008)	0.193 (0.008)
5 Year Dummies	—	Yes	Yes	Yes	Yes
9 Year-of-Birth Dummies	—	Yes	Yes	Yes	Yes
8 Region Dummies	—	Yes	Yes	Yes	Yes
Endogenous Regressors	—	Education	Veteran Status	Education & Veteran Status	NA
R-Squared	—	0.087	0.165	0.086	0.165
P-Value for Over-ID Test	—	0.531	0.525	0.511	NA

Notes: Sample size is 25,781. Data Set: March CPS, 1979, 1981-85, Special Lottery Match. Weekly Wage is in 1978 dollars. Instrumental variables are 130 lottery-dummy-by-year-of-birth interactions.

Table 4
2SLS and OLS Estimates of the Return to Schooling and Veteran Status

Independent Variables	2SLS (1)	2SLS (2)	2SLS (3)	2SLS (4)	OLS (5)
A. Dependent Variable: Log Annual Earnings, Men Born 1944-53, N=26,119					
Education	---	.078 (.002)	.091 (.024)	.091 (.024)	.079 (.002)
Veteran Status	-.065 (.089)	-.057 (.086)	-.056 (.086)	.024 (.086)	.023 (.011)
Endogenous Regressors	Veteran Status	Veteran Status	Education & Veteran Status	Education	---
B. Dependent Variable: Log Annual Earnings, Men Born 1950-53, N=11,291					
Education	---	.073 (.003)	.093 (.048)	.086 (.047)	.074 (.003)
Veteran Status	-.105 (.110)	-.111 (.108)	-.113 (.108)	-.036 (.027)	-.041 (.019)
Endogenous Regressors	Veteran Status	Veteran Status	Education & Veteran Status	Education	---
C. Dependent Variable: Log Weekly Earnings, Men Born 1950-53, N=11,132					
Education	---	.052 (.002)	.075 (.030)	.078 (.030)	.052 (.002)
Veteran Status	.036 (.070)	.032 (.068)	.029 (.068)	.004 (.018)	-.007 (.012)
Endogenous Regressors	Veteran Status	Veteran Status	Education & Veteran Status	Education	---

Notes: Exogenous variables are: veteran status dummy variable, 2 race dummies, a central city dummy, a balance of SMSA dummy, a marriage dummy, 5 year dummies, 9 year-of-birth dummies, and 8 region dummies. Instruments are 130 lottery-by-year-of-birth dummies.