Working Paper #251
Industrial Relations Section
Princeton University
April 1989

Lifetime Earnings and the Vietnam Era Draft Lottery:

Evidence from Social Security Administrative Records

by

Joshua D. Angrist[1]

Industrial Relations Section

Princeton University

April 1989

[1]Data collection for this project has been generously funded by the Industrial Relations Section at Princeton. Financial support of the Department of Economics, the Sloan Foundation and the Olin Foundation is also gratefully acknowledged. Numerous individuals and government agencies assisted in the production of the data sets used here. Part of this effort is described in a separate acknowledgement following the abstract. The views expressed are solely my own and do not reflect the views of any government agency or employee. This paper is part of my Princeton University Doctoral Dissertation. Special thanks go to my dissertation advisors Orley Ashenfelter, David Card, and Whitney Newey for their many helpful comments and suggestions.

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

Abstract

Estimates of the effect of veteran status on civilian earnings may be biased by the fact that certain types of men are more likely to serve in the armed forces. In this paper, an estimation strategy is employed that enables measurement of the effects of veteran status while controlling for differences in other personal characteristics related to earnings. The randomly assigned risk of induction generated by the Vietnam era draft lottery is used to construct instrumental variables that are correlated with earnings solely by virtue of their correlation with veteran status. Instrumental variables estimates tabulated from Social Security Administration records indicate that in the early 1980's the earnings of white veterans were approximately 15 percent less than nonveteran earnings. In contrast, there is no evidence that nonwhite veterans suffered any lasting reduction in earnings. In an attempt to explain the loss of earnings to white veterans, experience-earnings profiles are estimated jointly with time-varying veteran status coefficients. The estimates suggest that the effect of Vietnam era military service on white veterans is equivalent to a loss of two years of civilian labor market experience.

Joshua D. Angrist
Industrial Relations Section
Firestone Library
Princeton University
Princeton, NJ 08540

(609) 452-6374

Acknowledgements

The earnings data used in this study were prepared by the Office of Research, Statistics and International Policy (ORSIP) at the Social Security Administration's Baltimore Headquarters. The project was supervised by ORSIP director Warren Buckler and undertaken by data base managers Ada Enis and Cresston Smith. Programming and production of the final product was the work of Bea Matsui. Thanks also go to Bev Smith for her prompt response to my inquires about self-employment. Of particular importance was the help of the ORSIP staff during the Internal Revenue Service disclosure review.

The data set constructed from Defense Manpower Data Center (DMDC) Administrative Records is the work of Mike Dove and his staff of programmers at the DMDC's Monterey office. In addition to the main data set, Mr. Dove provided a number of related files, usually on short notice.

The version of the Survey of Income and Program Participation (SIPP) used here is the work of the Chief of Income Surveys at the Census Bureau, Chester Bowie, and the members of his staff. Mr. Bowie's patient efforts led to the construction of a data set that would be useful to me while still meeting Census Bureau disclosure requirements.

Linda Stalvey, Selective Service National Headquarters Director of Public Affairs, provided helpful background material on selective service in the Vietnam era.

I. Introduction

One-third of American men born between 1944 and 1953 served in the armed forces, the vast majority during what is now called the "Vietnam era".[1] As the men who came of age in the Vietnam era move into positions of responsibility, questions regarding the incidence and consequences of military service have become increasingly important. For example, the 1988 presidential election campaign included statements of support for veterans benefits and raised questions about the candidates' response to the draft (Wall Street Journal 1988a and 1988b, Newsweek 1988). Discussions of alternative means of military manpower procurement have also resurfaced (New York Times 1987).

A central issue in the debate over military manpower policy is the question of whether veterans are adequately compensated for their service. In particular, attention has focused on the costs imposed upon men who were drafted for service in Vietnam or who "volunteered" for service to avoid the draft. The desire to compensate those who served has had considerable impact on the political process: in 1989 the federal government plans to spend nearly 30 billion dollars on veterans benefits and services (Council of Economic Advisors 1988). Yet, in spite of this vast expenditure, it is far from clear whether veterans are truly worse off than nonveterans. A member of the Twentieth Century Fund Task Force on Policies Toward Veterans has argued that, "Within any age group, veterans have higher incomes, more education and lower unemployment rates than their nonveteran counterparts" (Taussig 1974, p. 51).

The subject of this paper is the measurement of long-term labor market consequences of military service in the Vietnam Era. Previous academic research

[1] The Vietnam Era officially includes the period between August 5, 1964 and May 7, 1975 (Veterans Administration 1981b). The one-third figure is tabulated from the March 1985 Current Population Survey.

on this topic, which generally finds that Vietnam Era veterans earn less than nonveterans, may not provide an accurate picture. This is because conventional estimates can be biased by the fact that certain types of men are more likely to serve in the armed forces. 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. In the econometric program evaluation literature, this is known as the problem of selection bias.[2] The research reported in this paper is distinguished from previous work by the use of the draft lottery to set up a natural experiment that enables measurement of the veteran effect without selection bias. In the experimental design generated by the draft lottery, inferences about the effects of military service can be made as if veterans status itself had been determined on the basis of random assignment.[3]

In the draft lotteries, priority for induction was determined by a Random Sequence Number (RSN) assigned to each date of birth in the cohort at risk of being drafted. The 1970 lottery involved men born in the years 1944-1950, the 1971 lottery involved men born in 1951, the 1972 lottery involved those born in 1952, and so on, through 1975. No one was inducted after 1972, however, and congressional authority to induct expired in July 1973. Over the course of the

[2]Studies that find Vietnam veterans earn less than nonveterans include J.J. Card (1983, Tables 4.1 and 4.4), DeTray (1982, Table 1), Rosen and Taubman (1982, Table 3), Schwartz (1986, Table 3) and Crane and Wise (1987, Tables 6.3 and 6.5). On the other hand, DeTray (1980) finds positive effects for whites and Veterans Administration (1981a) finds an overall positive effect. A candid assessment of the selection bias problem in this context is given by Crane and Wise, who note they were unable to use econometric selection models to generate robust estimates of the effects of military service on civilian earnings.

[3]This approach was first used in epidemiologic research by Hearst, Newman and Hulley (1986), who present lottery-based estimates of delayed effects of military service on mortality.

lottery year, men were called for induction by RSN. At some point an official RSN ceiling was determined by the Defense Department; only men with lottery numbers below the ceiling could be drafted. The 1970 induction ceiling was RSN 195, the 1971 induction ceiling was RSN 125 and the 1972 induction ceiling was RSN 95. Men with lottery numbers below the RSN ceiling are referred to here as "draft-eligible."

Much of the paper is devoted to showing how the differences in earnings by draft-eligibility status can be converted to estimates of the effects of Vietnam era military service. The effect of draft-eligibility on earnings and other results in the paper are tabulated from a special version of the Social Security Administration (SSA) Continuous Work History Sample (CWHS), a unique data set constructed from a one per cent sample of SSA administrative records. These data are described in detail in Section III, below. To motivate the statistical analysis, however, the effect of draft-eligibility on the social security taxable earnings of draft lottery participants is also described here in Figure 1. The figure shows cohort earnings profiles (in 1978 dollars)[4] for the four cohorts born between 1950 and 1953. For each cohort two lines are drawn, one for men with lottery numbers that made them draft-eligible and one for men with lottery numbers that exempted them from any risk of conscription.

The effect of draft-eligibility on civilian earnings is striking. For each of the cohorts born from 1950 to 1952, there appears to be no difference in annual earnings by draft-eligibility status until the year in which the cohort was at risk of induction. Subsequently, the earnings of draft-eligible white men fall below the earnings of white men who could not be drafted. The earnings of draft-eligible nonwhites also fall below the earnings of other nonwhites,

[4]The deflator used for all tabulations is the CPI on p. 313 of Council of Economic Advisors (1988).

however, the gap appears to close and become positive for some nonwhite cohorts in later years.

Estimation of the effects of military service from differences in earnings by draft-eligibility status is justified by the assumption that because draft lottery numbers were assigned randomly, they are correlated with earnings only by virtue of their correlation with veteran status. Thus, the reason draft-eligible men earn less than men who could not have been drafted is assumed to be solely because a higher fraction of draft-eligible men are veterans. Knowledge of the fraction of veterans in the draft-eligible and ineligible groups may then be used to convert differences in earnings by draft-eligibility status into estimates of the effects of military service. This set of assumptions may also be recognized as those that justify instrumental variables estimation. In the empirical work that follows, a variety of functions of randomly assigned lottery numbers are used as instruments for regressions of earnings on veteran status.

Results from the statistical analysis indicate that from 1981-84 the annual earnings of most white veterans continued to lag behind the earnings of nonveterans by approximately 15 percent, or $3,500 current dollars. On the other hand, there is no evidence of any long-term effects of military service on nonwhite earnings. Some attention is also given to a simple explanation for white veterans' earnings loss. The hypothesis that the loss of earnings to white veterans is a consequence of lost civilian labor market experience is tested by estimating experience-earnings profiles jointly with time-varying veteran status coefficients. A model that attributes the impact of veteran earnings solely to loss of experience is found to provide a good description of white earnings from 1975 to 1984. The estimates suggest that white Vietnam veterans lost two years of labor market experience and that their earnings may not be expected to catch up with nonveteran earnings in the near future.

The plan of the paper is as follows. In the next section, additional features of the draft lottery are described and econometric issues raised by the use of the lottery to measure the effects of military service are discussed. The first step of lottery-based estimation is simply to compare the earnings of draft-eligible men to the earnings of men who were not draft-eligible. Statistical results from this comparison, along with a discussion of the earnings data, are presented in Section III.

Section IV introduces data on the relationship between veteran status and draft lottery numbers and presents results from the conversion of earnings differences by draft-eligibility status into estimates of the effects of military service. Section V presents results from a more efficient instrumental variables estimation strategy and Section VI corroborates results from the Social Security data with evidence from the National Longitudinal Survey of Young Men (NLS) and the Survey of Income and Program Participation (SIPP). Section VII discusses theory and evidence on the loss-of-experience interpretation of the veteran penalty. Finally, Section VIII offers a summary and some directions for further research.

II. The Draft Lottery and Military Service

A. National Random Selection[5]

A few months before the year in which a cohort was slated for induction, lottery numbers were drawn on national television and the results published in the newspapers. In addition to the RSN, selective service registrants may also

---

[5] The material in this section draws heavily on Tarr (1981) and Selective Service System (1986).

have been ranked on the basis of the first letter of their last name. In practice, however, further selection seems to have been largely based on the screening process employed by local draft boards and by the elimination of those with deferments. Deferments were granted for a variety of educational, occupational or family reasons and men with deferments would retain the liability implied by the RSN attached to their date of birth upon expiration of the deferment.

Selection of individuals for induction from the draft-eligible, non-deferred "high priority pool" was based on a number of criteria, the most important of which were the pre-induction physical examination and the examination for mental aptitude. In 1970, for example, half of all registrants failed pre-induction examinations and 20 percent of those remaining were eliminated by physical inspections conducted at the time of induction (Selective Service System 1971, p. 5-6). Of course, the fact that the selection process for entry into the military was ultimately non-random does not imply that the priority for induction was not randomly assigned.

A important contribution of the draft to military manpower procurement was the production of "draft-motivated volunteers" who enlisted to avoid induction under less favorable circumstances (House Armed Services Committee 1970, Tarr 1981). The behavioral response to the draft lottery was such that even in cohorts from which there were few inductions, there are large differences in probabilities of enlistment by Random Sequence Number (Angrist 1989b). Because induction authority did not officially expire until the middle of 1973, this behavioral response can also be detected in enlistment rates for the 1953 cohort. Therefore, for purposes of comparison by draft-eligibility status, the "draft-eligibility ceiling" for the 1953 cohort is set at RSN 95, which is the highest lottery number called from the cohort inducted the preceding year.

The year 1970 was the last time nondeferred men over the age of 19 were drafted. Men born between 1944-49 continued to be at risk of induction in the 1970 lottery, however, most veterans from these cohorts had already entered the service by the time of the 1970 lottery. Those who managed to "hold out" until 1970 may not be representative of their entire cohort. For this reason the analysis below is restricted to a study of men who turned 19 in the year in which they were at risk of induction as a consequence of the lottery. This sample includes only men who were born between 1950 and 1953.

B. Using the Draft Lottery to Measure the Effects of Military Service

Measurement of the effects of military service is a program evaluation problem of the type considered by Ashenfelter and Card (1985) and Heckman and Robb (1985). As is typical in program evaluation research, interest here centers on parameters in a simple linear model for earnings.

Let individual i's earnings at time t be denoted by y_{it} and let s_i be an indicator of veteran status. Then we may write

$$(1) \qquad y_{cti} = \beta_c + \delta_t + s_i\alpha + u_{it} \qquad \text{for } t > k$$

where k is the time of entry into service, β_c is a cohort effect, δ_t is a period effect common to all cohorts and u_{it} is the regression error. α is the "treatment effect" which it is desired to estimate. If s_i is correlated with unobserved components of the earnings equation, then α will not be consistently estimated by Ordinary Least Squares (OLS). Correlation between s_i and u_{it} may arise either because armed forces eligibility criteria are correlated with earnings but unobserved by the econometrician, or because veterans are self-selected on the basis of unobserved characteristics.

The draft lottery facilitates estimation of (1) by providing instrumental variables that are correlated with s_i, yet, by virtue of their random assignment, 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 had a lottery number below the induction ceiling for his year of birth. 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:

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

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.

Intuitively, equation (2) simply adjusts earnings differences by draft-eligibility status for the fact that not all draft-eligible men actually served while some men who were not draft-eligible voluntarily enlisted. 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 veteran is responsible for differences in earnings by draft-eligibility status. This formula may also be recognized as an application of Wald's (1940) grouping method, where the data have been grouped by draft-eligibility status. The fact that Wald's method is a form of instrumental variables (IV) estimation was first pointed out by Durbin (1954).

Some of the estimates reported below are based on the strategy embodied in (2). From the analogy to instrumental variables, however, it should be apparent that one can do better than the simple Wald estimator. In particular, given information on veteran status and draft lottery numbers, the most informative instrument generated by the draft lottery is the relationship between RSN and the probability of military service. Alternately, if the underlying probability model is unknown, one may employ polynomial terms in the RSN as instruments (Heckman 1978, Kelejian 1971). This is the approach taken in Angrist (1989a) using NLS micro data.

The data analyzed here do not consist of individual earnings histories. Rather, the observations consist of summary statistics for cells defined by year of birth, year of earnings, race and groups of five consecutive lottery numbers. Nevertheless, relatively efficient instrumental variables procedures may also be used to analyze the grouped data. To see this, consider the following generalization of the instrument set used to generate the Wald estimator. Instead of a single indicator for draft-eligibility status, the new instrument matrix has columns that indicate groups of 5 consecutive lottery numbers for each race, cohort and year of earnings. There are 73 columns corresponding to a particular race, cohort and year: the first column indicates men with lottery numbers 1-5 and the 73rd column indicates men with lottery numbers 360-365.

Now consider the following grouped version of (1), where $\bar{y}_{ctj}$ is mean earnings at time t for members of cohort c with lottery numbers in group j and $\hat{p}_{cj}$ is the fraction of cohort c with lottery numbers in group j who served:

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

Generalized Least Squares (GLS) estimates of (3) may easily be shown to be the same as instrumental variables estimates of (1), where the instrument matrix is of the form described above. Furthermore, the quadratic form minimized by the GLS estimator is a Generalized Method of Moments (GMM) over-identification test statistic associated with the use of dummy variables as instruments (Angrist 1988).[6] The test statistic may also be viewed as a measure of the goodness of fit of the sample means to model (1).

Along with the Wald estimator, equation (3) provides the basic framework for inference used in this paper. Implementation of the estimation strategy is straightforward - most estimates are simply coefficients from Generalized Least Squares (GLS) regressions of mean social security earnings on estimates of $\hat{p}_{cj}$ derived from a variety of sources. Before turning to a detailed discussion of data and results, however, it is important to clarify what the reported parameter estimates represent. To that end, a number of identification problems are reviewed below.

C. Identification of Veteran Status Treatment Effects

1. Treatment Effect Heterogeneity

Model (1) is a highly stylized representation of the program evaluation problem, allowing for neither treatment effect heterogeneity nor covariates. A modest generalization is the random coefficients model:

$$y_{cti} = \beta_c + \delta_t + s_i\alpha_i + u_{it} \qquad (4)$$
$$\alpha_i = \alpha_0 + \varepsilon_i,$$

where α_i is the treatment effect experienced by individual i, with population mean equal to α_0.

The introduction of a random treatment effect highlights the fact that the effect of veteran status can only be estimated for those who served in the armed

[6] A general reference on GMM estimation and testing is Newey (1985).

forces. If veterans are more or less likely to benefit from military service than the rest of the population, the estimated treatment effect does not characterize the impact of veteran status on a random sample. Formally, when ε_i is assumed uncorrelated with the instruments, an instrumental variables estimator of (4) identifies (Heckman and Robb 1985)

$$\alpha^* = \alpha_0 + E(\varepsilon_i | s_i=1). \tag{5}$$

Related to the problem of treatment effect heterogeneity is the fact that not all Vietnam era accessions to the military were draft-induced. A substantial fraction of enlistments were made by "true volunteers" (Tarr 1981), that is, men who would have volunteered for service in the absence of draft. Suppose that the impact of military service on the earnings of true volunteers differs from the effect of military service on the earnings of draftees and men who enlisted because of the draft. Then use of functions of the draft lottery 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 individual i. Note that functions of draft lottery numbers will only be correlated with f_i. Therefore α_g is not identified and $\alpha_g g_i$ will become part of the regression error.[7]

---

[7]In the random coefficients model this problem may be thought of as arising because of correlation between α_i and the instruments. I thank Kevin M. Murphy for helpful discussions on this point.

2. The Absence of Covariates

An additional problem arises because the administrative records used in this study contain no information on covariates other than race and age. Lack of covariates 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.[8] The need for covariates, denoted by x_i, may be represented by replacing δ_t with $x_i\delta$ in (1). In this case (assuming the instruments are uncorrelated with α_i), instrumental variables estimates will identify

$$\tilde{\alpha} = [E(x_i|s_i=1)-E(x_i|s_i=0)]\delta + \alpha_0 + E(\varepsilon_i|s_i=1).$$

On the other hand, the fact that veteran status influences civilian earnings primarily through its influence on third variables might be of subordinate importance for issues related to veterans compensation.

---

[8]This is in spite of the use of a randomly assigned instrument. Although the instrument is orthogonal to u_{it}, covariates "caused" by s_i are correlated with the instrument. Suppose that education, z_i, is related to s_i by

$$z_i = s_i\delta_1 + \nu_{1i}.$$

Then valid instruments for s_i will usually be correlated with z_i as well.

s_i may be also correlated with components of the regression error, say a_i, so that

$$s_i = a_i\delta_0 + \nu_{0i}.$$

Valid instruments, although correlated with s_i, are not correlated with a_i. The intuition here is that a randomly assigned instrument is uncorrelated with any characteristic determined _prior_ to veteran status, but it is correlated with anything determined _by_ veteran status.

3. Earnings-modifying Draft Avoidance Behavior

Perhaps the most serious identification problem arises if the risk of induction affected earnings through some other mechanism 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 so as to avoid having to flunk students out of school. Baskir and Strauss (1978, p. 29) claim that Vietnam era college enrollment was "6-7 percent higher than normal because of the draft".

If earnings-modifying draft-avoidance behavior is correlated with lottery numbers, lottery-based instruments will be correlated with the regression error and estimates of the effects of military service constructed using the draft lottery may be biased. However, in a previous analysis of data with covariates (Angrist 1989a), specification tests provided no evidence of a relationship between lottery numbers and characteristics other than veteran status. Furthermore, in the present context, the over-identification test statistic associated with use of functions of the lottery as instruments may be expected to detect some forms of omitted variables misspecification.

III. Draft-Eligibility and Social Security Earnings

The empirical analysis begins with a discussion of the Social Security data set used to tabulate differences in earnings by draft-eligibility status and lottery number.

A. The Matched Continuous Work History Sample

The earnings data are taken from a version of the Social Security

Administrations' Continuous Work History Sample (CWHS). The CWHS, described in detail in Appendix C, is a one percent sample drawn from a sampling frame of all possible social security numbers. Included in the CWHS are two earnings series. The first, beginning in 1964 in the version used here, contains information on the earnings histories of individuals in employment covered by FICA (Social Security), and 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 what follows, I refer to the first series as FICA earnings and to the second as W-2 earnings.

The FICA earnings series has two main drawbacks. First, the lack of information on uncovered employment is a well-known and potentially troubling problem in research using Social Security Administrative (SSA) records. The earnings of an individual who did not work or was employed only in uncovered employment for the entire year appear as a zero in the CWHS. There is no way to distinguish between "true zeros" - caused by a year long spell of unemployment or withdrawal from the labor force - and zeros caused by uncovered employment or missing data. Particularly disturbing is the possibility that veterans employed in government are missing from the CWHS. Recent estimates by Cohany (1988) indicate that in 1987 21% of veterans were employed in the public sector, as compared to 12% of nonveterans. On the other hand, roughly 70% of state and local government workers and over half of all government workers were covered by Social Security by 1981 (Meyer 1985, Table 2.1 and Appendix C, below). In fact, by 1980, roughly 90% of all workers were engaged in Social Security covered employment (U.S. Dept. of Health and Human Services 1987, Table 4). Covered employment has also included members of the Armed Forces since 1956.

The second drawback of the FICA earnings series is that earnings from any single source are censored at the Social Security taxable maximum. That is, earnings from a single source in excess of the taxable maximum are recorded as being at the maximum. Other researchers working with SSA administrative records have found that this type of censoring can have a substantial impact on results (e.g., Rosen and Taubman 1982). It should also be noted that FICA earnings from more than one source are censored by source and men with multiple sources may have FICA earnings that exceed the FICA taxable maximum.

The W-2 earnings series originated in 1978 when the SSA began requiring employers to file employee earnings reports on an annual basis. Previously, employers had been required to make quarterly reports. The importance of the switch to annual reporting is that, under the new system, W-2 forms containing uncensored wage and salary earnings are submitted for workers in both covered and uncovered employment. The problem with the W-2 series is that there was a long period of transition from quarterly reporting to annual reporting during which the W-2 series were not reliable.[9]

The original CWHS data set does not contain information on exact date of birth. For the purposes of research using the draft lottery, this information was matched to the CWHS from other administrative sources (see Appendix C). SSA programmers then matched lottery numbers to exact dates of birth using the tables in Selective Service System (1969-73).

The matched version of the CWHS had to meet the Internal Revenue Service disclosure requirements imposed on all data collected for tax purposes. As a result, CWHS data on individual earnings could not be released. Instead, the SSA provided aggregate data for cells defined by year of earnings, year of birth,

---

[9]Thus, only in 1985 did SSA actuaries begin using their own W-2 series to construct the Average Monthly Wage series needed to index Social Security benefits. Previously, indexation had been based on Internal Revenue Service data (U.S Department of Health and Human Services 1987, Table I).

race and five consecutive lottery numbers. The aggregate data set contains means, variances, number of decedents as determined by benefit status, fraction at the taxable maximum, fraction above the taxable maximum, fraction zero and number in each cell.

Descriptive statistics for the CWHS are reported in Table 1. Because the analysis is restricted to men born between 1950-53, sample statistics are presented only for this group. The combined statistics for four cohorts were constructed by computing weighted averages of cohort means. Information for the oldest cohort is available back to 1964, however the earliest information in the table is for 1969. Unless otherwise noted, statistics in Table 1 and all subsequent tables refer to men with positive earnings.

The descriptive statistics indicate that after 1972 the fraction of men in the sample with zero FICA earnings varies roughly between 15 to 22 percent for whites and between 33 to 35 per cent for blacks. To evaluate these figures, note that the author's tabulations show that roughly 10% of white men in these cohorts have zero recorded wage and salary earnings in late 1970's Current Population Surveys (CPS). Suppose that of the 90% who work, 12% are in the uncovered sector so that only 79% of the cohort may be expected to have positive FICA wage and salary earnings. Adding an estimated 5% who only have FICA self-employment earnings implies that 16% should have zero FICA earnings of any type.[10] Thus, 14-18% zero for whites' 1973 to 1980 CWHS earnings 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

---

[10] Approximately 6.4 per cent of men born between 1944-53 had only self-employment earnings in 1984. (Figures from correspondence with Ms. Bev Smith, Social Security Administration Headquarters, Baltimore). Presumably this fraction is lower for the younger 1950-53 cohorts.

social security earnings.[11] Problems of undercoverage and filing delay may be especially severe for nonwhites. The 15% of CPS nonwhites with zero wage and salary earnings is not large enough to rationalize the approximately 34% of nonwhites with zero FICA earnings in the CWHS.

The fraction with FICA earnings at the taxable maximum is more variable than the fraction zero, ranging from 3 to 15 percent for whites and between 2 to 10 percent for blacks. The fraction of men with FICA earnings above the taxable maximum, usually around 1-2%, is reported in parentheses in the column containing the fraction exactly at the FICA limit. Note that although the fraction with W-2 earnings above the FICA taxable maximum is also recorded in the table, the W-2 data are not censored at the maximum.

W-2 and FICA earnings have roughly equal fractions at the FICA taxable maximum, suggesting that both variables are drawn from the same underlying distribution. However, problems with early years of the W-2 series are reflected in the sample statistics. For example, the standard deviation of whites' W-2 earnings in 1978 is 6 times as large as 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 to 1980. The W-2 series also contains a substantially higher fraction of zeros than does the FICA series. Part of this difference, however, is due to the inclusion of self-employment earnings in the FICA series. Also, there are generally some individuals with FICA taxable earnings but no federally taxable compensation (Millea and Kilss 1980). These caveats notwithstanding, the W-2 data do appear reasonable for the years 1981-84.

[11] This is the explanation offered by Card and Sullivan (1988) who also observe increasing fractions with zero earnings in the Continuous Longitudinal Manpower Survey (CLMS) data set derived from Social Security administrative records. Social Security Administration programmers have confirmed to me that long filing delays account for the large and increasing fraction of Social Security zeros in recent years.

B. Draft-Eligibility Treatment Effects

Figure 2 shows the time series of draft-eligibility treatment effects for FICA taxable earnings. In each panel of the figure, the four ruled horizonal lines indicate no difference in earnings by draft-eligibility status for each of the four cohorts. Differences in earnings by draft-eligibility status are plotted relative to the ruled horizontal lines.

As in Figure 1, it is apparent that for each cohort there is no difference during the years before the lottery in which the cohort was at risk of induction, whereas in subsequent years the earnings histories diverge. Draft-eligible white men born between 1950 and 1952 generally appear to earn less after 1973. The lack of a difference in earnings by draft-eligibility status before the lottery is to be expected if, in fact, lottery numbers were randomly assigned. It is also apparent that the loss of earnings to draft-eligible white men born between 1950-52 is largest during their military service. However, the earnings of draft-eligible white men from these cohorts continued to lag behind the earnings of draft-ineligible men through 1984.

The picture for nonwhites is less clear than that for whites. 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, the temporal pattern of earnings differences for nonwhites born in 1952 is similar to that for whites. The general impression for the three older cohorts of nonwhites is that, by 1984, the earnings of draft-eligible men have at least caught up with the earnings of draft-ineligible men.

Earnings of white men born in 1953 do not appear to differ by draft-eligibility status and the earnings of draft-eligible nonwhites born in 1953 generally exceed the earnings of nonwhites who were not draft-eligible. The difference between the effect of draft-eligibility on men born in 1953 and the three older cohorts might be explained by the transition to an All-Volunteer Force during 1973. The improvement in military pay and benefits during this period could have been enough to offset any disadvantage suffered by veterans (Cooper 1977, Chapter 2).

Graphical descriptions, of course, do not tell the whole story. Draft-eligibility treatment effects and their standard errors are presented in Tables 2, 3 and 4. Table 2 reports treatment effects for FICA earnings (corresponding to Figure 2) as well as effects for W-2 earnings. Table 3 reports the effect of draft-eligibility on the probability of being censored at the FICA taxable maximum and Table 4 reports the effect of draft-eligibility on the probability of having zero recorded earnings in the CWHS.

The statistics in Table 2 show that the loss in FICA earnings to draft-eligible white men born is sometimes statistically significant and amounts to 2-3% 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 are rarely larger than their standard errors.

Table 3 shows that draft-eligible white men not only earned less, they were also less likely to have earnings above the FICA taxable maximum. 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 earnings treatment effects. This result is important because, as shown below, when the effect of draft-eligibility on the probability of being censored has the same sign as the effect on earnings, results tabulated using censored earnings data tend to underestimate the true effect.

Finally, note that the figures in Table 4 indicate that draft-eligible whites were somewhat less likely to have zero FICA earnings during the years in which they were in the service. The data also suggest that nonwhites may be less likely to have had zero earnings in recent years. There is no statistically significant evidence, however, of any lasting effect of draft-eligibility on the probability of having zero earnings for either racial group.

IV. Wald Estimates of the Effects of Military Service

This section works through the steps required to calculate Wald estimates of the effects of veteran status using formula (2). The numerator of the formula consists of the draft-eligibility treatment effects tabulated in Table 2. The denominator is the effect of draft-eligibility on veteran status. Information on the relationship between draft-eligibility and the probability of veteran status is obtained from a combination of sources and statistical procedures, described below.

A. The Matched SIPP and DMDC Administrative Records

The first set of probabilities is estimated using a special version of the 1984 Survey of Income and Program Participation. The SIPP data, described in detail in Appendix D, were matched to an indicator of draft-eligibility status using date of birth information on the Census Bureau's in-house version of the SIPP file. Because of the small number of observations available for single-cohort statistics, each SIPP probability is actually the average for 3 consecutive cohorts. For example, SIPP estimates assigned to men born in 1951 are based on data for men born from 1950-52.

The second set of probabilities comes 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 non-prior-service accessions by cohort and lottery number from July 1970 through December 1973.

Estimates of overall cohort size are derived from the CWHS. Recall that the CWHS is a 1 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 form an estimate of the probability of being a veteran conditional on being draft-eligible, the number of draft-eligible accessions 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 an estimated binomial proportion.[12]

The columns labelled $\hat{p}^e$ and $\hat{p}^n$ in Table 5 show probabilities of veteran status tabulated using the SIPP and DMDC/CWHS procedures described above. The last column of Table 5, labelled $\hat{p}^e$-$\hat{p}^n$, shows the difference in the probability of military service by draft-eligibility status. Estimates range from a high of .16 in SIPP estimates for the 1950 white cohort, to a low of .01 in DMDC/CWHS estimates for the 1953 nonwhite cohort.

The SIPP and DMDC/CWHS based estimates are also compared in Table 5 to unconditional probabilities of Vietnam era veteran status estimated from the March 1984 Current Population Survey (CPS). As a consequence of the lack of DMDC data on accessions before July 1970, SIPP and CPS marginal probabilities of

---

[12] The formula used is $\surd[\hat{p}(1-\hat{p})/n_c]$ 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 = .21 with estimated variance (.21*.79)/5749.

service for most cohorts exceed the estimates constructed using the DMDC/CWHS procedure. Missing accessions data will only be a problem, however, if the probability of being missing varies by lottery number. This is clearly the case for calculations involving the 1950 cohort because the 1970 accessions of this cohort are highly correlated with draft-eligibility status. In fact, SIPP based estimates of the difference in probability of service by draft-eligibility status are much greater than DMDC/CWHS estimates for men born in 1950. Therefore, only the SIPP data will be used in calculations involving the 1950 cohort. Note that for men born between 1951 and 1953, the SIPP and DMDC/CWHS estimates of $\hat{p}^e-\hat{p}^n$ are reasonably close.

B. Wald Estimates

Wald estimates of the effect of military service on the 1981-84 earnings of white men born between 1950 and 1952 are presented in Table 6. To construct these estimates, DMDC/CWHS based estimates of the effects of draft-eligibility on veteran status are used in an application of formula (2) to men born in 1951 and 1952. Because of the effect of missing 1970 data on DMDC estimates for men born in 1950, the SIPP data are used for this cohort. Estimates of the effect of draft-eligibility on veteran status are close to .15 for all three cohorts. 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/.15 = 6 2/3.

Draft-eligibility treatment effects for three different earnings variables are shown in Table 6. Column (1) presents treatment effects for FICA earnings and column (3) presents treatment effects for W-2 earnings. The figures in both columns are copied directly from Table 2. In addition, column (2) of Table 6 reports treatment effects 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 more detail in the next subsection. Note that the adjusted treatment effects are usually bracketed by the unadjusted FICA and W-2 treatment effects. Therefore, only Wald estimates constructed from the adjusted data are reported in the table.

The Wald estimates, reported in column (5) of the table, indicate that white veterans suffered an earnings loss of roughly two thousand 1978 dollars or $3,500 current dollars. This is approximately 15 percent of annual W-2 compensation for white men between 1981 and 1984. Standard errors for the Wald estimates are derived from the t-statistics associated with the draft-eligibility treatment effects.[13] The similarity of coefficient estimates across cohorts and years suggests 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.

C. Non-parametric Correction for Censoring

This section briefly outlines the procedure used to adjust FICA taxable earnings for censoring at the taxable maximum. For economy of notation, the c,t subscripts are dropped and all cells indexed by j. Let the population means for each cell be denoted by

μ_j^0 = mean of uncensored earnings

μ_j^c = mean of CWHS (censored) earnings

---

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

μ_j^t = mean of earnings truncated at the taxable maximum

L_j = the applicable FICA taxable maximum

μ_j^ℓ = mean of earnings for men with CWHS earnings above the taxable maximum

$p_j^{x\ell}$ = the probability of having CWHS earnings exactly at the taxable maximum

p_j^ℓ = the probability of having CWHS earnings greater than the taxable maximum.

A key assumption used here is that the mean CWHS earnings of men with earnings above the taxable maximum are the same as population means above the taxable maximum. Then we may write

$$\mu_j^c = p_j^{x\ell} L_j + p_j^\ell \mu_j^\ell + [1 - p_j^{x\ell} - p_j^\ell]\mu_j^t$$

$$\mu_j^0 = (p_j^{x\ell} + p_j^\ell)\mu_j^\ell + [1 - p_j^{x\ell} - p_j^\ell]\mu_j^t$$

so that,

$$\mu_j^0 = \mu_j^c + p_j^{x\ell}(\mu_j^\ell - L_j).$$

The FICA earnings series is adjusted for censoring by using the above formula and estimates of μ_j^ℓ tabulated from March Current Population Surveys (CPS) for each year, race and cohort.[14] This procedure is justified if μ_j^ℓ does not vary by lottery number. In other words, military service is assumed to effect the earnings of those below the taxable maximum as well as the probability of being at or above the taxable maximum. But there is assumed to be no effect on

[14]Estimation of standard errors for the adjusted series is discussed in Appendix A. The combined CPS means above the taxable maximum for the four cohorts are also presented in the appendix. Note that the CPS means are themselves censored, although the CPS censoring thresholds are substantially larger than the FICA taxable maximums.

mean earnings conditional on earning more than the taxable maximum.

The relationship between censored and true mean earnings may be used to examine the consequences of CWHS censoring for treatment effects estimated using uncensored data. Suppose that cell j contains a sample of draft-eligible men for a given race, cohort and year and that cell k contains a corresponding 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 analog of

$$\mu_j^c - \mu_k^c = (\mu_j^0 - \mu_k^0) + (p_j^{x\ell} - p_k^{x\ell})(L - \mu_j^\ell) + p_k^{x\ell}(\mu_k^\ell - \mu_j^\ell).$$

Note that the proposed correction for censoring will adjust the treatment effects for $(p_j^{x\ell} - p_k^{x\ell})$, but that the third term in this decomposition, $p_k^{x\ell}(\mu_k^\ell - \mu_j^\ell)$, is taken to be zero.

Now assume that the effect of draft-eligibility on both $p^{x\ell}$ and μ^ℓ is of the same sign as the effect on μ^0. Then the treatment effect estimated from the censored data differs from the true treatment effect by terms that are opposite in sign from the true effect. The magnitude of the treatment effect is therefore underestimated in the censored data. Thus, it is not surprising that estimates tabulated using the adjusted data tend to be bracketed by estimates tabulated from the unadjusted FICA and uncensored W-2 data.

V. Efficient Estimation Strategies

The Wald estimator highlights the analogy between the draft lottery and a randomized "natural experiment" that may be used to study the consequences of military service. The main drawback of the Wald estimator, however, also derives from its simplicity. It is apparent from the large standard errors in Table 6 that to generate precise estimates of the effects of military service, the data must be combined in a more efficient estimation strategy. Grouped data regression models such as equation (3) give efficient estimators and provide a convenient framework for the control of period and cohort effects. The regression framework also exploits the fact that more is known about the sample than draft-eligibility status. In particular, for each race, cohort and year of earnings, 73 cell means are available to estimate the parameters of interest.[15]

Estimates based on equation (3) are tabulated by regressing CWHS cell means on conditional probabilities of veteran status in each five-number-cell ($\hat{p}_{cj}$). The $\hat{p}_{cj}$ for men born from 1951 - 1953 are estimated using the DMDC/CWHS procedure described in Section IV. The SIPP data are used to construct probabilities for the 1950 cohort. Because the SIPP sample is too small to allow estimation of a complete set of $\hat{p}_{cj}$ for all lottery number cells, SIPP-based probabilities are computed for only two cells, defined by draft-eligibility status. The CWHS earnings data for men born in 1950 are then aggregated correspondingly. Thus, the estimating sample for each race and year includes 73 cell means for each of the three cohorts born from 1951-53 plus 2 cell means for the 1950 cohort.

A graphical interpretation of equation (3) is presented in Figure 3. The figure shows the relationship between the probability of veteran status and mean W-2 compensation (in 1978 dollars) between 1981 and 1984. The plotted points consist of the average (over four years of earnings) residuals from a regressions of earnings and probabilities on period and cohort effects.[16] Thus, the slope of the ordinary least squares regression line drawn through the points corresponds

---

[15] In Angrist (1988), GLS on grouped data is shown to be the minimum variance linear combination of all pairwise Wald estimators that can be computed from any set of grouped observations.

[16] 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.

to an estimate of α. This slope is equal to -2384 with a standard error of 778. A further 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 variance by draft-eligibility status; draft-eligible men have somewhat less variable earnings.

As pointed out in Section II, 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 probabilities is complicated by the use of multiple samples. A general distribution theory for the optimal two-sample IV estimator is outlined in Angrist (1989c) and application to the current context is discussed in Appendix B. Assuming that the samples used to calculate mean earnings and the sample used to calculate $\hat{p}_{cj}$ are independent, the optimal estimator has a simple form that may be described using the following notation. 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 two-sample IV estimator chooses θ to minimize

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

which is also the GLS minimand for (3).[17]

[17] The estimator may also be motivated as an application of Optimal Minimum Distance (OMD) techniques such as those described by Chamberlain (1982). Ignoring period and cohort effects, OMD estimates for the current problem are tabulated by choosing α and p to minimize

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

The minimized value of $m(\theta)$ is an overidentification test statistic that tests the validity of the use of dummy variables as instruments. If some of these dummy variables are correlated with the error term in equation (1), then $m(\theta)$ will 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 7 presents two sets of 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, whereas Model 2 restricts estimated service effects to be the same across cohorts. Note that, as in Table 6, "adjusted FICA earnings" are FICA-taxable earnings adjusted for censoring at the taxable maximum using the procedure described in the previous section.

The results in Table 7 show that white veterans born from 1950-52 suffered an annual earnings loss of between $1,500 and $2,100 1978 dollars. These results are generally consistent with the Wald estimates reported in Table 6. Also as in Table 6, regression estimates for adjusted FICA earnings tend to be bracketed by the results for unadjusted FICA and W-2 earnings. All the estimates for whites in Model 2 are larger than twice their standard errors. In contrast, results for nonwhites show no evidence of a statistically significant earnings loss to veterans.

Degrees of freedom for the chi-square overidentification test statistics reported at the bottom of Table 7 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,

---

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

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.

The chi-square statistics reported in the table never take on values larger than their degrees of freedom, suggesting that the residuals in equation (1) are not correlated with lottery-based instruments. 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 consider only the omnibus goodness-of-fit test.

The difference between the test statistics for Models 1 2 is a chi-square test for the restriction of equal treatment effects. This set of restrictions 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 3, indicating that the estimated treatment effects are not statistically different. The results for 1953, however, do appear qualitatively different.

VI. Additional Evidence from the SIPP and NLS

Estimation using the CWHS requires information from a variety of sources and the use of somewhat unconventional procedures. To corroborate the CWHS results, estimates tabulated using SIPP and NLS micro data are also reported. Unlike the CWHS, both the SIPP and the NLS contain information on individual earnings and dates of birth. Thus, these data sets may be used to construct lottery-based estimates of the effects of military service using traditional instrumental variables methods.

A number of lottery-based instruments may be employed when working with SIPP and NLS micro data. For NLS respondents it is possible to determine individual lottery numbers. Therefore, the veteran effect in the NLS sample is estimated here (and in Angrist 1989a) using polynomial terms in the lottery number as well as an indicator of draft-eligibility as instruments.

The matched SIPP does not contain lottery information more detailed than draft-eligibility status. However, because the Selective Service failed to properly randomize the first draft-lottery, there is some correlation between month of birth and the probability of having a low lottery number for men born in 1950 (Fienberg 1971).[18] SIPP-based estimates are constructed using draft-eligibility status and a full set of month of birth dummies as instruments. Both SIPP and NLS equations include year of birth dummies.[19] The reported estimates were tabulated using Whites's (1982) efficient, heteroscedasticity-consistent procedure.

The NLS sample does not include men born after 1952 so both samples are restricted to men born between 1950 and 1952. An unavoidable difference between the two samples is that SIPP earnings variables are for 1983-84 whereas NLS earnings variables are for 1981. The SIPP data are described in detail in Appendix D and the NLS data are described in the footnotes to the table and in the appendix to Angrist (1989a). Briefly, the SIPP equations are for monthly

---

[18]The draft lotteries were based on physical randomization. In the 1970 lottery, capsules containing dates of birth for each month were dumped into a barrel and then mixed. As it turns out, RSN's drawn from the barrel generally gave lower numbers to the men whose dates of birth fell in the months dumped in the barrel last. Subsequent lotteries used a two-barrel system and were designed in consultation with statistician John Tukey (Tarr 1981).

[19]The draft-eligibility ceiling for the 1950 cohort was 195, for the 1951 cohort, 125, and for the 1952 cohort, 95. Because older cohorts were more likely to be draft-eligible, it is important to include cohort dummies or some other variable to control for age when combining cohorts and using draft-eligibility status as an instrument.

wage and salary earnings on the respondent's main job and for a measure of average hourly earnings. The NLS equations are for annual wage and salary earnings and for the hourly rate of pay. Equations for both data sets are estimated in logarithms.

Results from the SIPP and NLS, reported in Table 8, are broadly consistent with results from the CWHS. Although imprecise, estimated earnings effects for whites are negative in both data sets. Estimates for nonwhites are also negative, though less than half the size of their standard errors.

The magnitude of the point estimates for whites is also similar to magnitudes found using the Social Security data. NLS estimates suggest an annual earnings penalty of approximately $2,700 1978 dollars. Multiplying results for SIPP monthly earnings by 11.5 gives an implied annual earnings penalty of approximately $2,500 1978 dollars. The magnitude of estimated veteran coefficients from regressions in levels, not reported here, tends to be a few hundred dollars lower for whites. NLS based estimates from levels regressions for nonwhite earnings are positive but imprecise.

Estimates for hourly earnings have lower standard errors than those for annual and monthly earnings, showing a negative effect for whites and a positive effect for nonwhites. This pattern of racial differences is consistent with the NLS results reported in Angrist (1989a). Along with coefficient estimates, GMM over-identification test statistics for instrumental variables estimates are also reported. Not surprisingly, given the large standard errors of estimated coefficients, the GMM tests provide no evidence of misspecification.

VII. Explaining the Veteran Earnings Loss

A simple explanation for the reduction in white veterans' earnings is that white veterans lost civilian labor market experience for which their military experience is only a partial substitute.[20] A particularly attractive feature of this explanation is that, given a functional form for the experience-earnings profile, the loss-of-experience model has testable implications.

One commonly encountered functional form for experience-earnings profiles is the log-quadratic (e.g., Mincer 1974). When the log of individual earnings is a quadratic function of experience, average log earnings for cohort c at time t may be written

$$\bar{y}_{ct} = \delta_t + \beta_0 x_c + \gamma x_c^2 + \bar{u}_{ct} \tag{6}$$

where $\bar{y}_{ct}$ now denotes average log earnings, δ_t is a period effect common to all cohorts, β_0 and γ are parameters and x_c is experience of the cohort at time t, taken here to be equal to t - (c + 18). Now suppose that the effect of military service is equivalent to a loss of ℓ years of experience. Then average log earnings for members of cohort c at time t in cell j may be shown to be

$$\bar{y}_{ctj} = \delta_t + \beta_0 x_c + \gamma x_c^2 - [\beta_0 \ell - \gamma \ell^2]\hat{p}_{cj} - [2\gamma\ell](\hat{p}_{cj} * x_c) + \bar{u}_{ct}.^{21} \tag{7}$$

A generalization of model (7) allows the linear term in (6) 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

---

[20]Alternative explanations include DeTray's (1982) screening hypothesis and detrimental consequences of military service such as "post-traumatic stress syndrome" for combat veterans (Hearst, Newman and Hulley 1986).

[21]This is derived by working with the equivalent model for individual earnings and using the fact that $E(s_i|c,j) = E(s_i^2|c,j) = p_{cj}$. Note that the time-varying intercepts in model (7), and model (8), below, differ from those in equation (6).

(8) $\bar{y}_{ctj} = \delta_t + \beta_0 x_c + \gamma x_c^2 - [\beta_0 \ell - \gamma \ell^2 + \beta_1 \ell]\hat{p}_{cj} - [2\gamma\ell - \beta_1](\hat{p}_{cj} * x_c) + \bar{u}_{ct}.$

Both earnings function models have the same reduced form in terms of unrestricted regression coefficients:

$$\bar{y}_{ctj} = \delta_t + \beta_0 x_c + \gamma x_c^2 + \pi_1 \hat{p}_{cj} + \pi_2 (\hat{p}_{cj} * x_c) + \bar{u}_{ct}. \tag{9}$$

Thus, model (8) contains four structural and four reduced form parameters, excluding the time-varying intercept. The structural parameters are β_0, β_1, γ and ℓ, and the reduced form parameters are β_0, γ, π_1 and π_2. Model (7) imposes one testable restriction on the reduced form by setting $\beta_1 = 0$.[22]

One implication of the earnings function models is that veteran earnings eventually overtake nonveteran earnings. To see this, note that the reduced form veteran effect is $\pi_1 + \pi_2 x_c$, which equals zero when $x_c = -\pi_1/\pi_2$. Standard human capital regression results suggest that β_0 is a positive number and that γ is a small negative number, in which case π_1 is negative and π_2 is relatively small and positive. The veteran effect is therefore large and negative for those with few years of experience, but increases linearly toward zero and becomes positive as experience increases. It should be noted, however, that for model (7) the overtaking age (where $x_c = -\pi_1/\pi_2$) must be past the peak of the earnings profile, (7).[23] On the other hand, in model (8) the overtaking age may be either before or after the peak of the profile.

---

[22] A third model is derived by letting both γ and β vary with veteran status. This model leads to a reduced form similar to (9), with the only modification being the addition of a linear term of the form $\pi_3(\hat{p}_{cj} * x_c^2)$. In the empirical work, however, no evidence was found that such a term belongs in the earnings function reduced form.

[23] $-\pi_1/\pi_2$ is equal to $x^* + \ell/2$, where x^* is the profile peak. I thank Hank Farber for pointing this out. Thanks also go to John Pencavel, who discovered an error in an earlier version of this section.

Equations (7) and (8) are estimated here using the CWHS data on white veterans earnings. It also interesting, however, to examine the implications of the model using estimates from earnings functions reported in the literature. As an example, I have chosen parameter estimates from Rosen and Taubman (1982, Table 3), who work with Social Security data similar to the CWHS.[24]

Fitting a log-quadratic earnings function, Rosen and Taubman estimate β_0 to be .073 and γ to be -.001 . The implied veteran effects derived from model (7) and assuming 2 or 3 year losses of experience are tabulated below:

			Proportional earnings loss			
			6	7	8	9
ℓ	π_1	π_2	years after discharge from service			
3	-.228	.006	-.192	-.186	-.180	-.174
2	-.150	.004	-.126	-.122	-.118	-.114

It is apparent that the Rosen and Taubman parameters generate veteran effects of roughly the same magnitude as the veteran effects estimated here for white men in the early 1980's. The veteran overtaking age implied by π_1 and π_2 is equal to 56 (38 years of experience) assuming 3 years lost experience and 55.5 years of age assuming 2 years lost experience. Finally, it may be noted that Rosen and Taubman estimate a Vietnam veteran effect of -.197 in a variation on the regression they use to generate the experience profile.

Table 9 shows results from Non-Linear Least Squares (NLLS) estimation of (7), (8) and (9) using the real FICA earnings of men born from 1950 to 1952. The weighting matrix used in NLLS estimation is derived in a manner similar to that

[24]The Rosen-Taubman sample is an exact match of CPS data to SSA and Internal revenue Service administrative records and contains the 1951-76 FICA taxable earnings of white males aged 18 to 65 throughout the sample period.

used to construct the estimates in Table 7. Details of the optimally weighted estimation procedure are discussed in Appendix B. Earnings functions are commonly fit in logs, and so the dependent variable is taken to be the log of cell mean earnings. The log of the mean is not the same as the mean of the log, however, the matched CWHS does not contain the mean of log earnings. To the degree that earnings are log-normally distributed, use of the log of the mean will provide a reasonable approximation.

The sample used to construct the estimates in Table 9 begins in the fifth year after the lottery in which members of the cohort participated. 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).[25] Evidence from Table 7 suggests that veteran effects are essentially zero for the 1953 cohort and so this cohort was excluded from the estimation.

Estimates of model (7) are presented in column (1) of Table 9 and of model (8) in column (2). The chi-square statistic in column (3) is for the overall goodness-of-fit of the reduced form, while 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 and the estimated loss-of-growth, though negative, is not statistically significant. The overidentification test statistic again takes on a value less than its degrees of freedom, indicating that model (7) is not at odds with the data.

The estimated loss of experience derived from model (7) is approximately 2 years. This is somewhat low relative to the median length of service and suggests that military experience may be a partial substitute for civilian

[25] The sample range begins in 1975 for men born in 1950, in 1976 for men born in 1951 and in 1977 for men born in 1952.

experience. The age at which veteran earnings overtake nonveteran earnings is derived from the unrestricted reduced form coefficients and shown in column (3). Reduced form estimates imply that veteran earnings overtake nonveteran earnings around age 50 with a standard error of 16.[26]

VIII. Summary and Conclusions

Estimation of the effects of veteran status is complicated by the fact that the earnings of veterans and nonveterans differ for many reasons besides the military service of the veterans. In this paper, an instrumental variables estimation strategy based on the draft lottery is used to measure the effects of Vietnam era veteran status while controlling for differences in characteristics other than veteran status. Data from three sources - Social Security administrative records, the Defense Manpower Data Center and the Survey of Income and Program Participation - are combined in lottery-based estimation of the impact of Vietnam era veteran status on a time series of Social Security earnings. The Social Security results are also corroborated with SIPP and NLS micro data.

Results of the lottery-based estimation strategy indicate that almost ten years after their discharge from service, white veterans who served at the close of the Vietnam era appear to be earning substantially less than nonveterans. The annual earnings loss to white veterans is estimated to be on the order of $3,500 current dollars, or roughly 15 percent of yearly wage and salary earnings in the

[26]Delta method standard errors for the overtaking age 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 -.00019 and the square roots of σ_1 and σ_2 appear in Table 9.

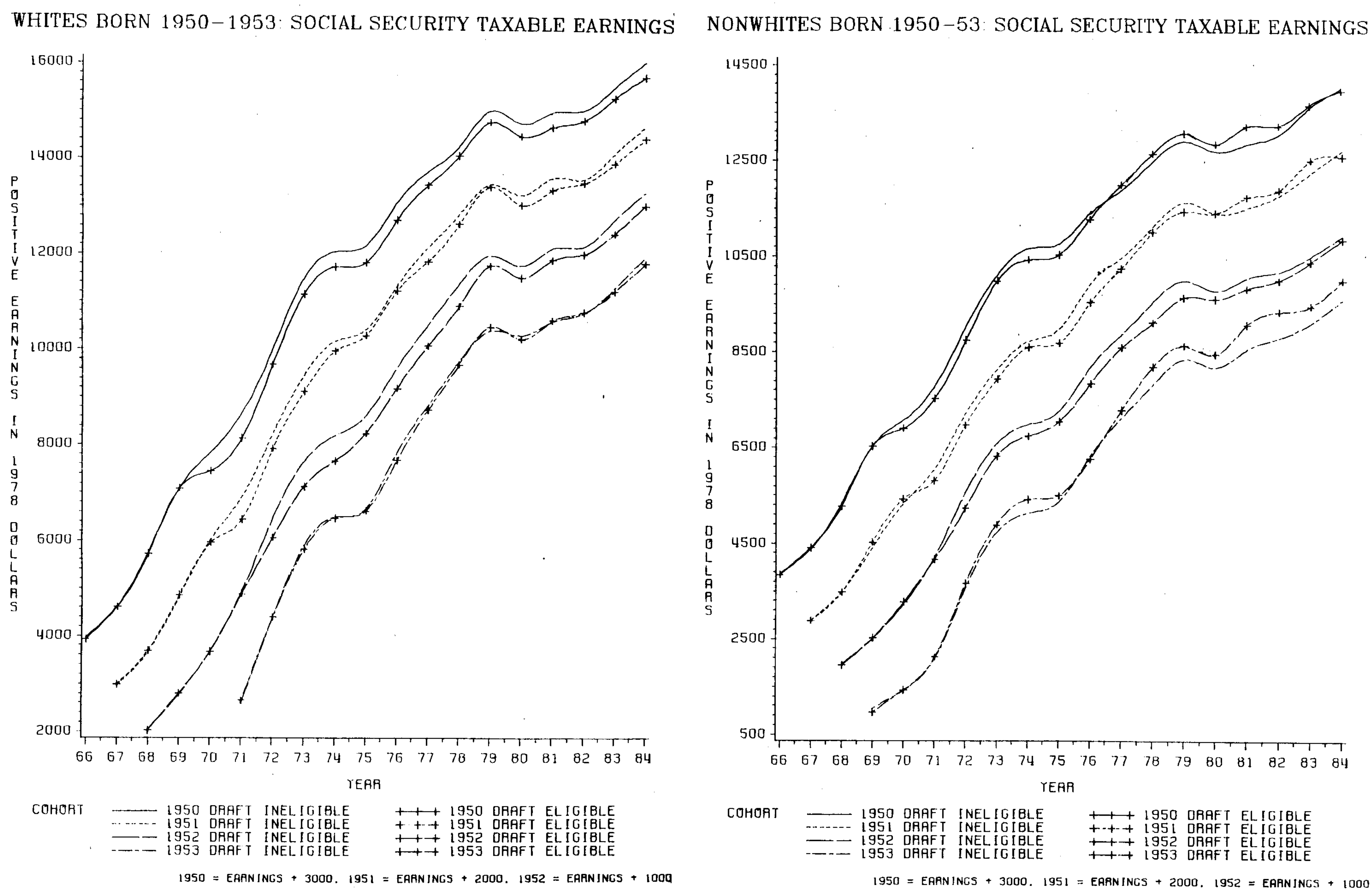
early 1980's. In contrast, the estimated veteran effects for nonwhites are not statistically significant.

A simple explanation for the loss of earnings to white veterans is also proposed: white veterans earn less because their military experience is only a partial substitute for the civilian labor market experience they lost while in the armed forces. This theory is tested by estimating experience-earnings profiles jointly with the effect of veteran status. Goodness-of-fit tests suggest that loss of civilian labor market experience provides an adequate explanation for the reduction in the earnings of white veterans. Estimated experience profiles imply white veterans suffered an earnings loss equivalent to two years of lost civilian labor market experience. Furthermore, reduced form equations predict that the earnings of white veterans will not overtake nonveteran earnings in the near future.

It should also be noted that some previous estimates of the effect of Vietnam era veteran status on the earnings of white men do not appear far off the mark. Estimates close to those reported here are those of Rosen and Taubman, who find a 19% annual earnings loss to Vietnam Era veterans and Crane and Wise (1987), who find an 11 percent reduction in 1979 weekly earnings. On the other hand, Berger and Hirsch (1983) find essentially no effect on 1977 weekly earnings and Angrist (1989) shows that OLS estimates of the effects of veteran status on hourly wages different substantially from lottery-based estimates.

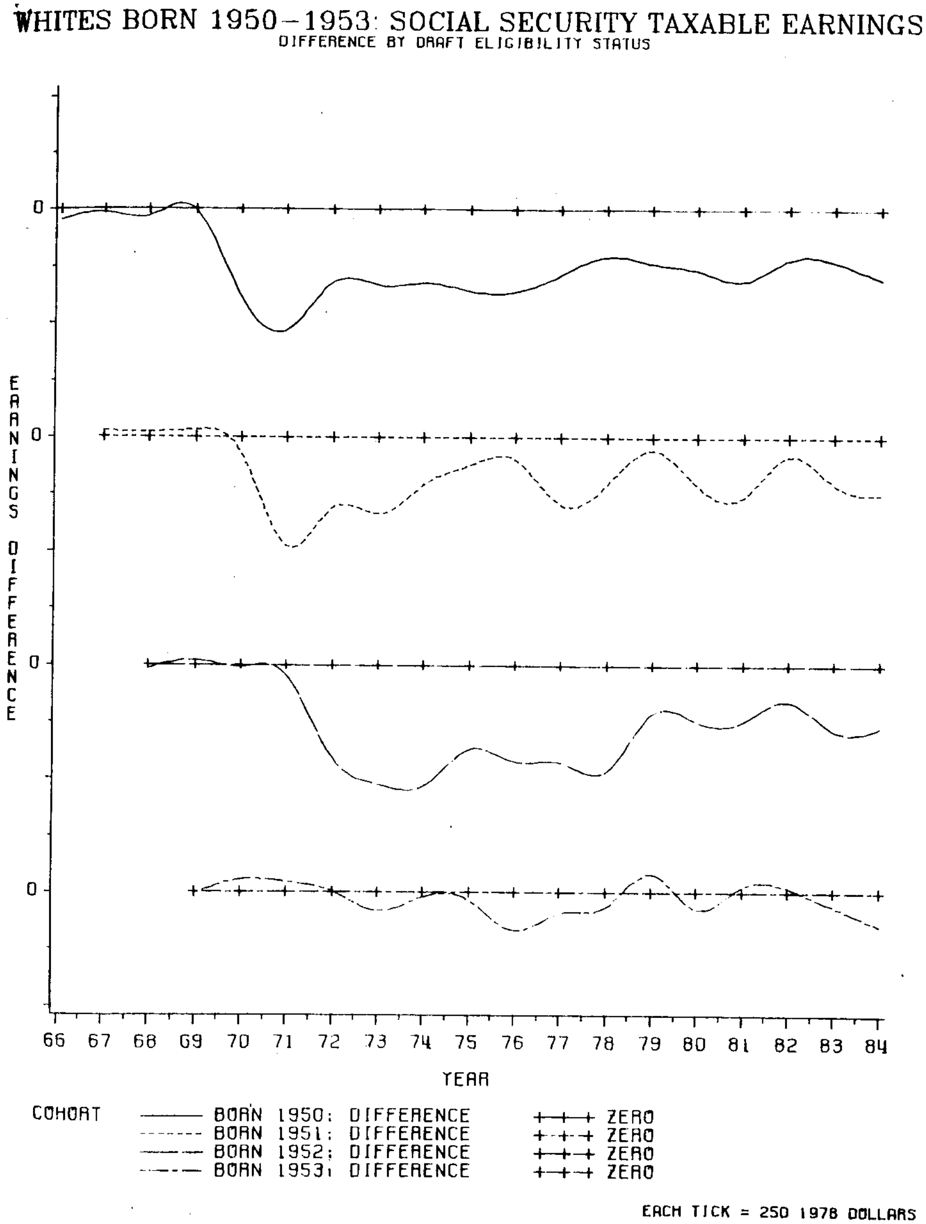
The estimation strategy used here involves a large number of small steps that must be carefully explained in a complete account of results and methods. To keep the paper to a manageable size, some issues have necessarily been left unexplored. Perhaps foremost among these is the question of alternatives to the loss of experience interpretation of the reduction in white veteran earnings. Interesting alternatives might be based on the notion that veteran status is a screening device (DeTray 1982) and on cohort size effects such as those discussed by Smith and 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 interesting avenue for future research is the use of CWHS information on decedents to extend the pioneering research of Hearst, Newman and Hulley (1986) on the effect of draft-eligibility on civilian mortality.

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 (e.g., 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, a strong case can be made that these results are actually 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 that 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.



WHITES BORN 1950–1953: SOCIAL SECURITY TAXABLE EARNINGS
POSITIVE EARNINGS IN 1978 DOLLARS
YEAR
COHORT
1950 DRAFT INELIGIBLE
1951 DRAFT INELIGIBLE
1952 DRAFT INELIGIBLE
1953 DRAFT INELIGIBLE
1950 DRAFT ELIGIBLE
1951 DRAFT ELIGIBLE
1952 DRAFT ELIGIBLE
1953 DRAFT ELIGIBLE
1950 = EARNINGS + 3000. 1951 = EARNINGS + 2000. 1952 = EARNINGS + 1000
NONWHITES BORN 1950–53: SOCIAL SECURITY TAXABLE EARNINGS
POSITIVE EARNINGS IN 1978 DOLLARS
YEAR
COHORT
1950 DRAFT INELIGIBLE
1951 DRAFT INELIGIBLE
1952 DRAFT INELIGIBLE
1953 DRAFT INELIGIBLE
1950 DRAFT ELIGIBLE
1951 DRAFT ELIGIBLE
1952 DRAFT ELIGIBLE
1953 DRAFT ELIGIBLE
1950 = EARNINGS + 3000. 1951 = EARNINGS + 2000. 1952 = EARNINGS + 1000

FIGURE 1


WHITES BORN 1950–1953: SOCIAL SECURITY TAXABLE EARNINGS
DIFFERENCE BY DRAFT ELIGIBILITY STATUS
EARNINGS DIFFERENCE
0
0
0
0
66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84
YEAR
COHORT
BORN 1950: DIFFERENCE
BORN 1951: DIFFERENCE
BORN 1952: DIFFERENCE
BORN 1953: DIFFERENCE
ZERO
ZERO
ZERO
ZERO
EACH TICK = 250 1978 DOLLARS

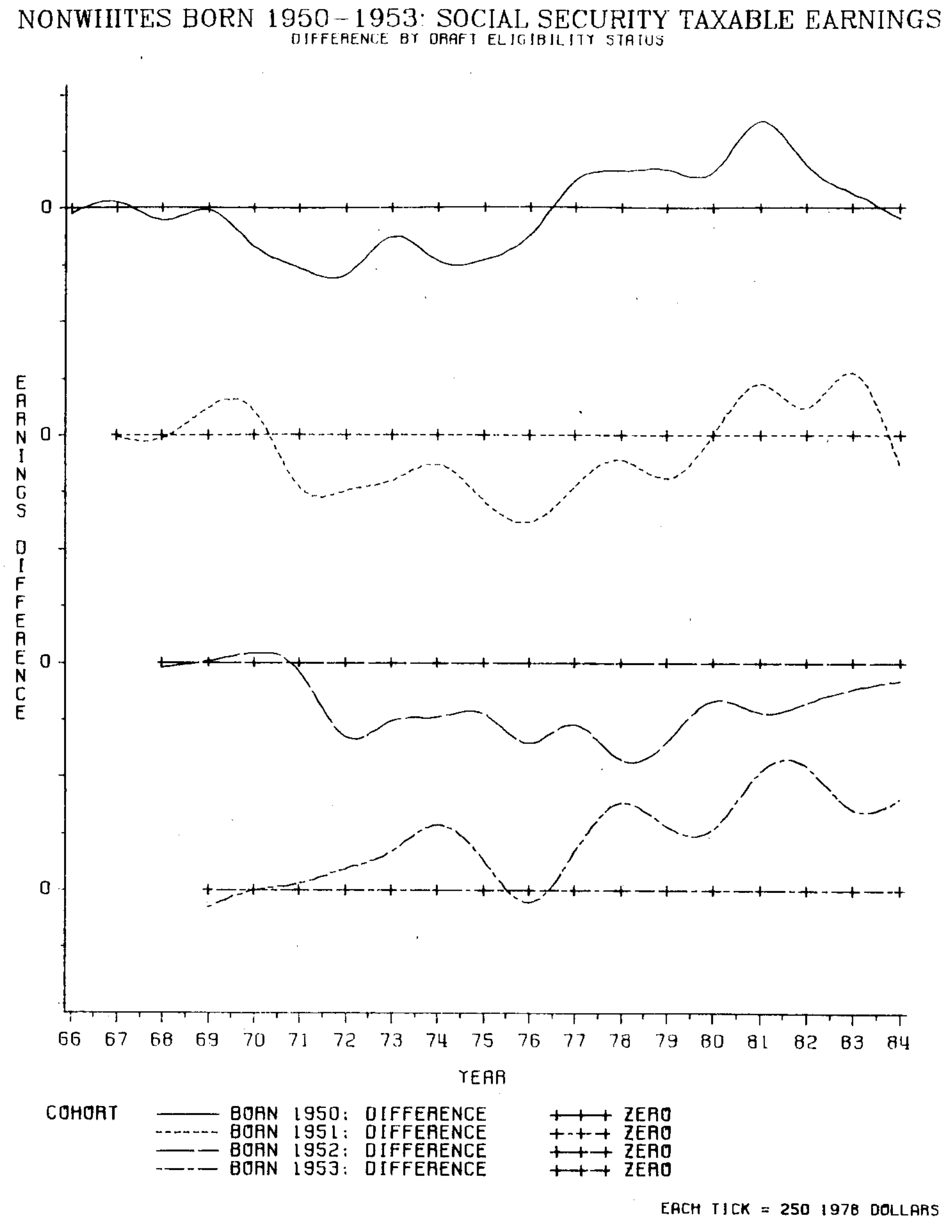
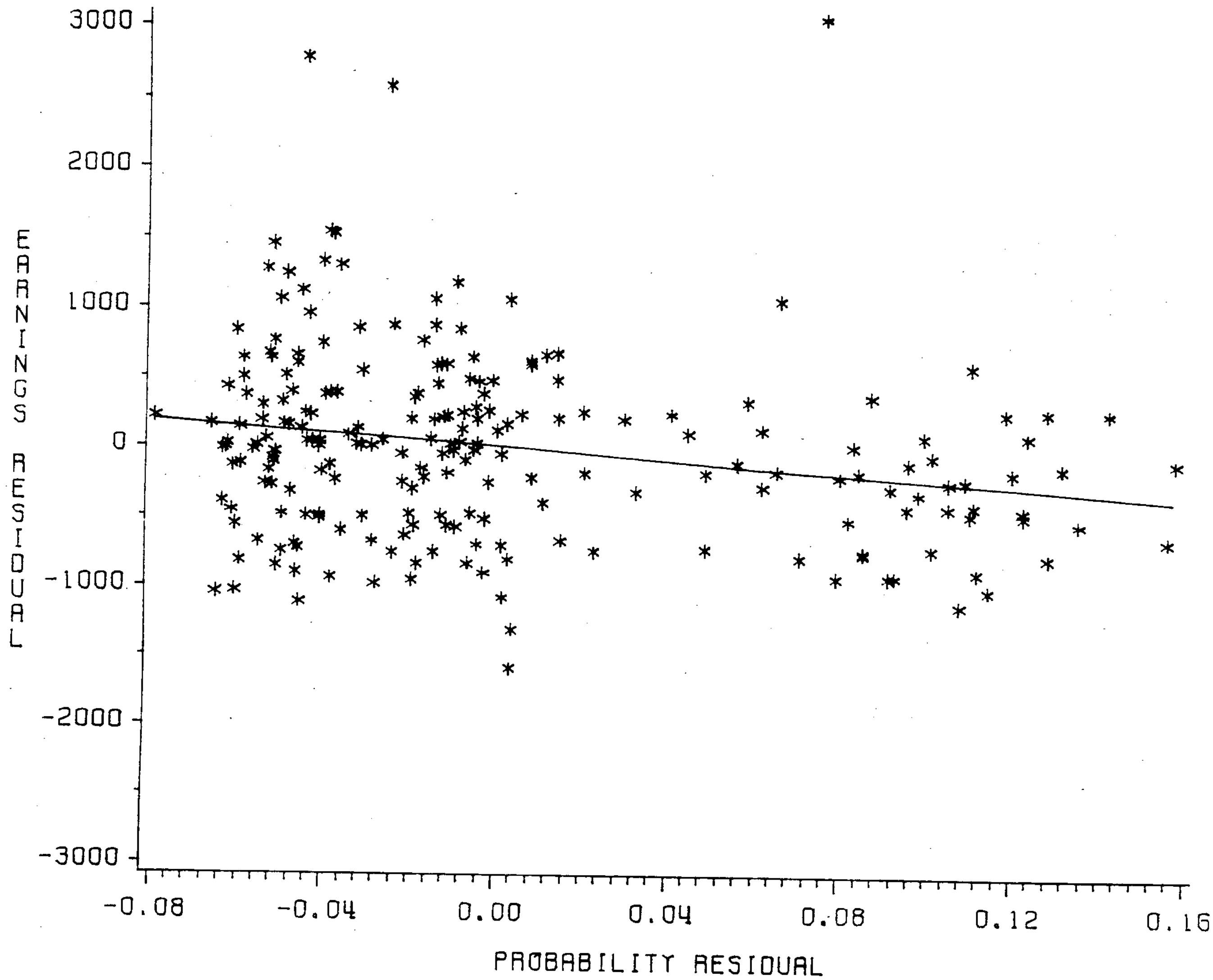

NONWHITES BORN 1950–1953: SOCIAL SECURITY TAXABLE EARNINGS
DIFFERENCE BY DRAFT ELIGIBILITY STATUS
EARNINGS DIFFERENCE
0
0
0
0
66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84
YEAR
COHORT
BORN 1950: DIFFERENCE
BORN 1951: DIFFERENCE
BORN 1952: DIFFERENCE
BORN 1953: DIFFERENCE
ZERO
ZERO
ZERO
ZERO
EACH TICK = 250 1978 DOLLARS

FIGURE 2

WHITES BORN 1950-53, TOTAL W-2 COMPENSATION 1981-84



MEAN EARNINGS AND PROBABILITY OF VETERAN STATUS BY LOTTERY NUMBER
AVERAGE RESIDUALS FROM REGRESSION ON PERIOD AND COHORT EFFECTS

FIGURE 3

Table 1

FICA Taxable (w/s and s/e) Earnings and Total W-2 (w/s) compensation
White Men Born 1950-53

Race	Year	N	Taxable Maximum	FICA Earnings	FICA at Limit[1,2]	FICA Zeros	W-2 Earnings[3]	W-2 at Limit[1]	W-2 Zeros[1]
White	69	68,407	7800	1,473 (1,457)	.003	.321			
	70	68,339	7800	1,977 (1,825)	.010	.241			
	71	68,244	7800	2,581 (2,233)	.030(.01)	.194			
	72	68,154	9000	3,614 (2,802)	.040(.02)	.157			
	73	68,053	10800	4,738 (3,460)	.046(.02)	.141			
	74	67,966	13200	5,727 (4,170)	.034(.01)	.145			
	75	67,882	14100	6,459 (4,843)	.053(.02)	.169			
	76	67,794	15300	7,698 (5,548)	.078(.02)	.167			
	77	67,691	16500	8,974 (6,206)	.107(.03)	.163			
	78	67,598	17700	10,441 (7,050)	.149(.04)	.165	15,435 (91,032)	.167	.244
	79	67,503	22900	12,388 (8,455)	.092(.02)	.166	14,786 (65,359)	.097	.261
	80	67,413	25900	13,769 (9,678)	.092(.02)	.175	14,561 (28,180)	.096	.272
	81	67,316	29700	15,641 (11,129)	.089(.02)	.183	16,363 (15,295)	.092	.272
	82	67,265	32400	16,743 (12,371)	.092(.02)	.200	17,907 (16,298)	.093	.282
	83	67,190	35700	18,046 (13,621)	.089(.02)	.210	19,595 (22,470)	.089	.275
	84	67,114	37800	19,717 (14,883)	.103(.02)	.219	21,595 (20,856)	.101	.281

Standard deviations of earnings in parentheses.

Table 1 (cont.): Nonwhites

Race	Year	N	Taxable Maximum	FICA Earnings	FICA at Limit[1,2]	FICA Zeros	W-2 Earnings[3]	W-2 at Limit[1]	W-2 Zeros[1]
Nonwhite	69	21,514	7800	1,306 (1,471)	.002	.464			
	70	21,501	7800	1,711 (1,865)	.007	.417			
	71	21,475	7800	2,197 (2,275)	.020(.01)	.388			
	72	21,443	9000	3,072 (2,922)	.025(.01)	.352			
	73	21,412	10800	3,992 (3,609)	.029(.01)	.330			
	74	21,380	13200	4,802 (4,359)	.021(.01)	.336			
	75	21,339	14100	5,404 (5,015)	.031(.01)	.359			
	76	21,310	15300	6,453 (5,836)	.049(.01)	.350			
	77	21,275	16500	7,510 (6,616)	.071(.02)	.341			
	78	21,243	17700	8,751 (7,541)	.096(.02)	.338	13,439 (63,722)	.116	.391
	79	21,200	22900	10,262 (8,938)	.061(.01)	.329	13,581 (63,244)	.067	.393
	80	21,167	25900	11,405 (10,147)	.063(.01)	.331	11,716 (13,895)	.067	.397
	81	21,130	29700	12,986 (11,628)	.063(.01)	.335	13,421 (14,644)	.063	.388
	82	21,109	32400	14,045 (12,849)	.067(.01)	.354	14,983 (16,112)	.065	.408
	83	21,077	35700	15,101 (14,167)	.067(.01)	.358	16,271 (19,089)	.065	.396
	84	21,042	37800	16,391 (15,237)	.077(.01)	.353	17,905 (20,784)	.073	.387

Standard deviations of earnings in parentheses.

Notes to Table 1:

(1) Fractions at limit are fractions of non zero earnings at or above FICA taxable maximum. Fractions zero are fractions of all observations in cell.

(2) 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. Figures in parentheses are fraction above taxable maximum.

(3) 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.

(4) Sample statistics are weighted averages of cells by race, year of birth, and 5 consecutive lottery numbers. Average cell sizes: 230 (whites) and 70 (nonwhites).

Table 2: Draft-Eligibility Treatment Effects for Earnings, Whites

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

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

Table 3: Draft-Eligibility Treatment Effects
Probability of Having Earnings at the Taxable Maximum, Whites

	FICA Taxable Earnings				Total W-2 Compensation			
Year	1950	1951	1952	1953	1950	1951	1952	1953
66	0.0000 (0.0000)							
67	0.0003 (0.0002)	0.0000 (0.0000)						
68	0.0002 (0.0004)	-0.0002 (0.0001)	0.0009 (0.0007)					
69	-0.0004 (0.0017)	0.0011 (0.0007)	0.0007 (0.0005)	0.0000 (0.0000)				
70	-0.0043 (0.0027)	-0.0028 (0.0018)	0.0001 (0.0007)	0.0001 (0.0003)				
71	-0.0170 (0.0043)	-0.0169 (0.0031)	-0.0015 (0.0023)	-0.0002 (0.0008)				
72	-0.0189 (0.0047)	-0.0144 (0.0035)	-0.0146 (0.0026)	-0.0004 (0.0018)				
73	-0.0216 (0.0049)	-0.0101 (0.0037)	-0.0122 (0.0028)	-0.0014 (0.0024)				
74	-0.0101 (0.0042)	-0.0038 (0.0034)	-0.0052 (0.0026)	-0.0021 (0.0021)				
75	-0.0116 (0.0050)	-0.0001 (0.0044)	-0.0072 (0.0034)	0.0001 (0.0028)				
76	-0.0147 (0.0056)	-0.0044 (0.0052)	-0.0097 (0.0043)	-0.0022 (0.0037)				
77	-0.0105 (0.0063)	-0.0085 (0.0057)	-0.0163 (0.0052)	0.0019 (0.0046)				
78	-0.0205 (0.0070)	-0.0111 (0.0066)	-0.0078 (0.0062)	0.0027 (0.0057)	-0.0160 (0.0077)	-0.0096 (0.0072)	-0.0007 (0.0070)	0.0043 (0.0063)
79	-0.0069 (0.0058)	-0.0000 (0.0053)	-0.0025 (0.0050)	-0.0028 (0.0045)	-0.0065 (0.0064)	0.0007 (0.0058)	-0.0026 (0.0055)	-0.0012 (0.0049)
80	-0.0035 (0.0058)	-0.0029 (0.0054)	-0.0099 (0.0050)	-0.0026 (0.0045)	-0.0064 (0.0063)	-0.0030 (0.0059)	-0.0098 (0.0054)	-0.0020 (0.0049)
81	-0.0044 (0.0056)	-0.0142 (0.0053)	-0.0121 (0.0050)	-0.0022 (0.0045)	-0.0058 (0.0061)	-0.0145 (0.0057)	-0.0115 (0.0054)	-0.0037 (0.0048)
82	-0.0063 (0.0058)	-0.0161 (0.0054)	-0.0082 (0.0052)	-0.0046 (0.0046)	0.0000 (0.0061)	-0.0203 (0.0057)	-0.0071 (0.0056)	-0.0037 (0.0050)
83	-0.0127 (0.0057)	-0.0155 (0.0053)	-0.0054 (0.0051)	0.0013 (0.0048)	-0.0058 (0.0060)	-0.0154 (0.0056)	-0.0058 (0.0054)	0.0067 (0.0051)
84	-0.0106 (0.0060)	-0.0173 (0.0057)	-0.0132 (0.0055)	0.0004 (0.0053)	-0.0093 (0.0062)	-0.0171 (0.0059)	-0.0159 (0.0056)	0.0024 (0.0055)

Table 3 (cont.): Nonwhites

Year	FICA Taxable Earnings 1950	1951	1952	1953	Total W-2 Compensation 1950	1951	1952	1953
66	0.0000 (0.0000)							
67	-0.0007 (0.0007)	0.0000 (0.0000)						
68	0.0005 (0.0005)	0.0000 (0.0000)	0.0000 (0.0000)					
69	-0.0001 (0.0027)	-0.0009 (0.0006)	-0.0004 (0.0004)	0.0000 (0.0000)				
70	-0.0033 (0.0042)	0.0016 (0.0034)	0.0009 (0.0014)	0.0000 (0.0000)				
71	-0.0083 (0.0067)	-0.0124 (0.0051)	-0.0020 (0.0033)	-0.0018 (0.0009)				
72	-0.0105 (0.0075)	-0.0126 (0.0051)	-0.0042 (0.0043)	0.0024 (0.0028)				
73	0.0001 (0.0078)	-0.0073 (0.0054)	-0.0045 (0.0048)	0.0041 (0.0039)				
74	0.0021 (0.0066)	-0.0081 (0.0049)	-0.0057 (0.0039)	-0.0049 (0.0026)				
75	0.0100 (0.0076)	-0.0035 (0.0065)	-0.0061 (0.0049)	-0.0043 (0.0041)				
76	-0.0047 (0.0090)	-0.0087 (0.0080)	-0.0069 (0.0070)	0.0064 (0.0062)				
77	-0.0070 (0.0104)	-0.0074 (0.0095)	-0.0151 (0.0080)	0.0109 (0.0080)				
78	0.0135 (0.0113)	-0.0062 (0.0109)	-0.0154 (0.0100)	0.0104 (0.0097)	0.0044 (0.0127)	0.0001 (0.0122)	-0.0194 (0.0117)	0.0078 (0.0112)
79	0.0067 (0.0091)	-0.0010 (0.0089)	-0.0123 (0.0080)	0.0167 (0.0080)	0.0113 (0.0101)	0.0074 (0.0097)	-0.0167 (0.0090)	0.0109 (0.0085)
80	-0.0033 (0.0093)	-0.0037 (0.0089)	-0.0094 (0.0082)	0.0158 (0.0084)	-0.0041 (0.0099)	-0.0027 (0.0099)	-0.0109 (0.0090)	0.0098 (0.0087)
81	0.0129 (0.0094)	0.0042 (0.0089)	-0.0077 (0.0081)	0.0131 (0.0086)	0.0094 (0.0099)	0.0065 (0.0095)	-0.0047 (0.0087)	0.0165 (0.0088)
82	0.0094 (0.0095)	0.0047 (0.0097)	-0.0165 (0.0083)	0.0133 (0.0088)	0.0101 (0.0098)	0.0103 (0.0102)	-0.0104 (0.0086)	0.0103 (0.0091)
83	0.0082 (0.0094)	0.0001 (0.0097)	-0.0149 (0.0086)	0.0157 (0.0089)	0.0078 (0.0096)	0.0063 (0.0099)	-0.0095 (0.0088)	0.0104 (0.0089)
84	-0.0015 (0.0099)	-0.0086 (0.0102)	0.0011 (0.0098)	0.0266 (0.0098)	-0.0050 (0.0100)	-0.0032 (0.0102)	0.0084 (0.0101)	0.0214 (0.0097)

Table 4: Draft-Eligibility Treatment Effects
Probability of have Zero Reported Earnings, Whites

	FICA Taxable Earnings				Total W-2 Compensation			
Year	1950	1951	1952	1953	1950	1951	1952	1953
66	0.0449							
66	(0.0076)							
67	0.0297	-0.0005						
67	(0.0076)	(0.0079)						
68	0.0153	-0.0137	0.0044					
68	(0.0066)	(0.0077)	(0.0084)					
69	0.0007	-0.0123	0.0073	0.0002				
69	(0.0059)	(0.0065)	(0.0081)	(0.0084)				
70	-0.0067	-0.0161	0.0054	0.0079				
70	(0.0059)	(0.0061)	(0.0072)	(0.0082)				
71	-0.0114	-0.0171	-0.0033	0.0005				
71	(0.0059)	(0.0060)	(0.0065)	(0.0072)				
72	-0.0076	-0.0180	-0.0107	0.0010				
72	(0.0057)	(0.0058)	(0.0060)	(0.0062)				
73	-0.0081	-0.0154	-0.0156	-0.0099				
73	(0.0057)	(0.0055)	(0.0056)	(0.0058)				
74	-0.0094	-0.0175	-0.0178	0.0006				
74	(0.0058)	(0.0056)	(0.0056)	(0.0059)				
75	-0.0060	-0.0152	-0.0071	-0.0012				
75	(0.0061)	(0.0061)	(0.0062)	(0.0062)				
76	0.0000	-0.0076	-0.0078	0.0023				
76	(0.0061)	(0.0061)	(0.0062)	(0.0061)				
77	0.0048	-0.0091	-0.0047	-0.0000				
77	(0.0061)	(0.0060)	(0.0062)	(0.0060)				
78	0.0039	-0.0091	-0.0080	-0.0054	-0.0038	-0.0141	-0.0111	-0.0051
78	(0.0061)	(0.0060)	(0.0061)	(0.0061)	(0.0070)	(0.0070)	(0.0072)	(0.0070)
79	0.0009	-0.0048	0.0025	-0.0028	-0.0052	-0.0119	-0.0012	-0.0123
79	(0.0062)	(0.0060)	(0.0062)	(0.0061)	(0.0072)	(0.0071)	(0.0074)	(0.0072)
80	-0.0070	-0.0102	0.0031	-0.0036	-0.0070	-0.0155	-0.0054	-0.0053
80	(0.0063)	(0.0061)	(0.0064)	(0.0063)	(0.0072)	(0.0072)	(0.0075)	(0.0073)
81	-0.0038	-0.0030	0.0034	-0.0029	-0.0004	-0.0099	-0.0075	-0.0105
81	(0.0063)	(0.0063)	(0.0066)	(0.0063)	(0.0072)	(0.0072)	(0.0076)	(0.0073)
82	0.0060	0.0027	0.0065	-0.0023	0.0044	-0.0041	0.0023	0.0024
82	(0.0065)	(0.0065)	(0.0068)	(0.0066)	(0.0073)	(0.0073)	(0.0077)	(0.0075)
83	0.0008	-0.0041	0.0048	0.0019	-0.0035	-0.0178	-0.0021	-0.0020
83	(0.0066)	(0.0066)	(0.0069)	(0.0068)	(0.0072)	(0.0072)	(0.0076)	(0.0074)
84	0.0026	0.0037	0.0028	-0.0094	-0.0011	-0.0074	-0.0011	-0.0137
84	(0.0068)	(0.0068)	(0.0070)	(0.0068)	(0.0073)	(0.0073)	(0.0077)	(0.0074)

Table 4 (cont.): Nonwhites

Year	FICA Taxable Earnings 1950	1951	1952	1953	Total W-2 Compensation 1950	1951	1952	1953
66 66	0.0133 (0.0124)							
67 67	0.0224 (0.0135)	0.0093 (0.0134)						
68 68	0.0335 (0.0132)	0.0002 (0.0145)	0.0195 (0.0145)					
69 69	0.0137 (0.0128)	-0.0025 (0.0140)	-0.0002 (0.0156)	-0.0132 (0.0149)				
70 70	0.0063 (0.0129)	0.0070 (0.0139)	0.0103 (0.0155)	-0.0157 (0.0155)				
71 71	-0.0057 (0.0129)	0.0026 (0.0139)	0.0239 (0.0154)	-0.0042 (0.0154)				
72 72	-0.0016 (0.0128)	-0.0033 (0.0136)	-0.0037 (0.0150)	0.0033 (0.0151)				
73 73	0.0093 (0.0126)	0.0004 (0.0135)	0.0028 (0.0149)	0.0060 (0.0147)				
74 74	0.0017 (0.0127)	0.0135 (0.0137)	0.0049 (0.0149)	0.0087 (0.0148)				
75 75	-0.0044 (0.0129)	-0.0003 (0.0138)	-0.0076 (0.0151)	-0.0061 (0.0150)				
76 76	0.0061 (0.0129)	-0.0089 (0.0138)	-0.0098 (0.0150)	-0.0158 (0.0148)				
77 77	-0.0011 (0.0128)	0.0064 (0.0137)	0.0033 (0.0150)	-0.0078 (0.0147)				
78 78	0.0056 (0.0129)	0.0034 (0.0136)	0.0056 (0.0150)	0.0036 (0.0148)	0.0002 (0.0133)	-0.0124 (0.0142)	0.0075 (0.0154)	0.0046 (0.0152)
79 79	-0.0025 (0.0128)	0.0005 (0.0136)	-0.0104 (0.0148)	-0.0065 (0.0147)	-0.0031 (0.0133)	-0.0171 (0.0142)	-0.0165 (0.0153)	0.0078 (0.0152)
80 80	-0.0049 (0.0129)	-0.0070 (0.0136)	-0.0101 (0.0148)	-0.0036 (0.0147)	-0.0104 (0.0134)	-0.0143 (0.0143)	-0.0103 (0.0153)	-0.0119 (0.0152)
81 81	0.0149 (0.0129)	0.0063 (0.0136)	-0.0209 (0.0148)	-0.0013 (0.0148)	0.0112 (0.0134)	0.0006 (0.0142)	-0.0263 (0.0152)	-0.0063 (0.0152)
82 82	-0.0043 (0.0131)	-0.0022 (0.0138)	-0.0171 (0.0150)	-0.0088 (0.0150)	0.0084 (0.0135)	-0.0085 (0.0143)	-0.0155 (0.0154)	-0.0052 (0.0154)
83 83	-0.0061 (0.0131)	0.0084 (0.0140)	-0.0091 (0.0151)	-0.0225 (0.0149)	-0.0022 (0.0134)	-0.0065 (0.0143)	-0.0118 (0.0154)	-0.0190 (0.0152)
84 84	0.0005 (0.0131)	-0.0137 (0.0139)	-0.0248 (0.0149)	-0.0268 (0.0149)	-0.0008 (0.0134)	-0.0241 (0.0142)	-0.0248 (0.0153)	-0.0251 (0.0151)

Table 5
Veteran Status and Draft-eligibility, Whites

Data Set	Cohort	Sample	P(Vet)	$[\hat{p}^e]$ P(Vet|elig)	$[\hat{p}^n]$ P(vet|inelig)	$[\hat{p}^e-\hat{p}^n]$ Difference
CPS (84)	50	1056	0.2761 (.0113)			
	51	1078	0.2191 (.0100)			
	52	1192	0.1790 (.0092)			
	53	1143	0.1309 (.0082)			
SIPP (84)	50	351	0.2673 (.0140)	0.3527 (.0325)	0.1933 (.0233)	0.1594 (.0400)
	51	359	0.1973 (.0127)	0.2831 (.0390)	0.1468 (.0180)	0.1362 (.0429)
	52	336	0.1554 (.0114)	0.2310 (.0473)	0.1257 (.0146)	0.1053 (.0495)
	53	390	0.1298 (.0106)	0.1581 (.0339)	0.1153 (.0152)	0.0427 (.0372)
DMDC/CWHS	50	16119	0.0633 (.0019)	0.0936 (.0032)	0.0279 (.0019)	0.0657 (.0037)
	51	16768	0.1176 (.0025)	0.2071 (.0053)	0.0708 (.0024)	0.1362 (.0059)
	52	17703	0.1515 (.0027)	0.2683 (.0065)	0.1102 (.0027)	0.1581 (.0071)
	53	17749	0.1343 (.0026)	0.1548 (.0053)	0.1268 (.0029)	0.0280 (.0060)

Standard errors in parentheses.

Notes to Table 5 (continued bottom of next page):

CPS (84) March 1984 Mare-Winship CPS extract.
Probabilities for service in the Vietnam Era (1964-75).

SIPP (84) Wave I, Panel I (1984) Survey of Income and Program Participation.
Probabilities for service in Vietnam Era.

DMDC/CWHS Defense Manpower Data Center Administrative Records' information on accessions, combined with information on cohort size from the 1% Social Security Adminstration Current Work History Sample of SSA administrative records.
Probabilities =
(# DMDC accessions July 1970-December 1973)/(CWHS cohort size in 1970).

Table 5 (cont.): Nonwhites

Data Set	Cohort	Sample	P(Vet)	$[\hat{p}^e]$ P(Vet|elig)	$[\hat{p}^n]$ P(vet|inelig)	$[\hat{p}^e-\hat{p}^n]$ Difference
CPS (84)	50	142	0.2255 (.0280)			
	51	150	0.1695 (.0256)			
	52	125	0.1344 (.0241)			
	53	135	0.1484 (.0253)			
SIPP (84)	50	70	0.1625 (.0292)	0.1957 (.0699)	0.1354 (.0491)	0.0603 (.0854)
	51	63	0.1703 (.0292)	0.2014 (.0827)	0.1514 (.0448)	0.0500 (.0940)
	52	52	0.1332 (.0275)	0.1449 (.1040)	0.1287 (.0373)	0.0161 (.1105)
	53	55	0.1749 (.0305)	0.1980 (.0865)	0.1612 (.0470)	0.0367 (.0984)
DMDC/CWHS	50	5447	0.0417 (.0027)	0.0548 (.0042)	0.0271 (.0032)	0.0276 (.0053)
	51	5258	0.0794 (.0037)	0.1173 (.0076)	0.0599 (.0040)	0.0574 (.0086)
	52	5493	0.0953 (.0040)	0.1439 (.0095)	0.0794 (.0042)	0.0644 (.0104)
	53	5303	0.0925 (.0040)	0.0984 (.0079)	0.0904 (.0046)	0.0080 (.0092)

Standard errors in parentheses.

Notes to Table 5 (cont.):

CPS (84) Weighted by CPS sampling weight.

SIPP (84) Weighted by SIPP sampling weight and smoothed over 3 cohorts. For example, estimates for 1953 are a weighted average of 1952, 53 and 1954.

DMDC/CWHS Standard errors calculated as $\hat{p}(1-\hat{p})/n_c$ where $\hat{p}$ is the estimated probability and n_c is 100 times the number in the social security cohort. For probabilities conditional on eligibility status, cohort size is calculated assuming lottery numbers have a uniform distribution in the population.

Table 6
The Effect of Military Service on the Earnings of White Men Born 1950-52

Year of Birth	Year	Current $ Treatment Effects for: FICA Earnings (1)	Adjusted FICA Earnings (2)	Total W-2 Earnings (3)	Prob Effect (4)	1978 $ Service Effect (5)
50	81	-435.8	-487.8	-589.6	0.159	-2,195.8
		(210.5)	(237.6)	(299.4)	(.040)	(1,069.5)
	82	-320.2	-396.1	-305.5		-1,678.3
		(235.8)	(281.7)	(345.4)		(1,193.6)
	83	-349.5	-450.1	-512.9		-1,795.6
		(261.6)	(302.0)	(441.2)		(1,204.8)
	84	-484.3	-638.7	-1,143.3		-2,517.7
		(286.8)	(336.5)	(492.2)		(1,326.5)
51	81	-358.3	-428.7	-71.6	0.136	-2,261.3
		(203.6)	(224.5)	(423.4)	(.006)	(1,184.2)
	82	-117.3	-278.5	-72.7		-1,386.6
		(229.1)	(264.1)	(372.1)		(1,312.1)
	83	-314.0	-452.2	-896.5		-2,181.8
		(253.2)	(289.2)	(426.3)		(1,395.3)
	84	-398.4	-573.3	-809.1		-2,647.9
		(279.2)	(331.1)	(380.9)		(1,529.2)
52	81	-342.8	-392.6	-440.5	0.158	-1,782.5
		(206.8)	(228.6)	(265.0)	(.007)	(1,037.9)
	82	-235.1	-255.2	-514.7		-1,091.3
		(232.3)	(264.5)	(296.5)		(1,131.1)
	83	-437.7	-500.0	-915.7		-2,076.5
		(257.5)	(294.7)	(395.2)		(1,223.8)
	84	-436.0	-560.0	-767.2		-2,226.3
		(281.9)	(330.1)	(376.0)		(1,312.3)

Notes: Standard errors in parentheses.
Columns (1) and (3) are taken from Table 2.
Column (2) is the draft-eligibility treatment effect on earnings adjusted for censoring at the FICA taxable maximum as described in Appendix A.
Column (4) is the SIPP based estimate of the effect of draft-eligibility on veteran status for men born 1950 and the DMDC/CWHS based estimate for men born 51-52, taken from Table (5). Column (5) is the effect of military service on civilian earnings implied by columns (2) and (4).

Table 7: Combined Service Effects, Whites
Optimal Instrumental Variables Estimates

	1981-84 Earnings in 1978 Dollars		
Cohort	FICA Taxable Earnings	Adjusted FICA Earnings	Total W-2 Compensation
Model 1			
1950	-1709.2 (946.8)	-2093.7 (1108.8)	-1895.0 (1333.1)
1951	-1457.1 (959.3)	-1983.7 (1036.1)	-2431.4 (1152.1)
1952	-1724.0 (863.1)	-1943.0 (927.2)	-2058.7 (1001.9)
1953	1223.8 (3232.1)	900.7 (3505.3)	-488.6 (3936.0)
$\chi^2(873)$	578.3	630.3	569.5
R^2	.986	.983	.978
Model 2			
1950-53	-1562.9 (521.8)	-1920.4 (575.9)	-2094.5 (646.3)
$\chi^2(876)$	579.1	631.0	569.7
R^2	.986	.983	.978

Table 7 (cont.): Nonwhites

Cohort	1981-84 Earnings in 1978 Dollars		
	FICA Taxable Earnings	Adjusted FICA Earnings	Total W-2 Compensation
Model 1			
1950	3893.7 (5358.5)	3891.9 (6244.5)	5711.8 (7206.0)
1951	-891.3 (4397.1)	-333.4 (4664.2)	2609.0 (4894.6)
1952	-3182.9 (3997.4)	-3457.7 (4195.2)	-3068.0 (4229.2)
1953	-5928.3 (10296.3)	-8571.4 (10697.1)	-6325.8 (11410.6)
$\chi^2(873)$	616.7	681.7	693.6
R^2	.919	.906	.886
Model 2			
1950-53	-643.3 (2407.5)	-999.7 (2602.5)	366.7 (2734.2)
$\chi^2(876)$	618.4	683.4	695.6
R^2	.931	.918	.901

Table 8: Evidence From the SIPP and NLS
Men Born 1950-52

Variable	Data Set	Race	Sample	Mean	Coefficient	GMM Test (dof)
Monthly Earnings	SIPP	White	778	1173.7 (24.9)	-.184 (.224)	6.5 (11)
		Nonwhite	132	938.7 (50.8)	-.269 (.637)	7.6 (11)
Annual Earnings	NLS	White	576	13237.4 (292.9)	-.208 (.377)	4.8 (3)
		Nonwhite	198	9504.2 (401.6)	-.187 (.555)	1.4 (3)
Average Hourly Earnings	SIPP	White	816	6.50 (.14)	-.278 (.196)	8.7 (11)
		Nonwhite	139	5.93 (.69)	.792 (.667)	9.1 (11)
Hourly Rate of Pay	NLS	White	541	6.99 (.37)	-.394 (.241)	3.4 (3)
		Nonwhite	194	5.16 (.19)	.168 (.300)	1.7 (3)

Notes: White (1982) efficient instrumental variables estimates; heteroscedasticity consistent standard errors in parentheses. Means are in 1978 dollars.

SIPP is Wave I, Panel I (1983/84) of the Survey of Income and Program Participation. SIPP instruments are draft-eligibility indicator, 11 month-of-birth dummies. SIPP monthly earnings are total wage and salary earnings in month preceding interview month. Average hourly earnings is hourly wage if paid hourly, else monthly earnings divided by weeks worked last month times hours per week.

NLS is the 1981 Wave of the National Longitudinal Survey of Young Men. NLS instruments are draft-eligibility indicator, cubic in draft lottery number. NLS hourly rate of pay is hourly wage if reported, else rate of pay per time unit of pay divided by hours worked in time unit of rate of pay. NLS veteran status is for service before 1976.

Table 9: Earnings Function Models for the Veteran Effect, Whites Born 1950-52

Parameter	Restricted: Loss of Experience (1)	Restricted: Loss of Experience, Reduced Growth Rate (2)	Unrestricted: OLS Estimates (3)
experience slope, β_0	.1022 (0.007)	.1016 (0.007)	.1016 (0.007)
experience squared, γ	-0.0027 (0.0003)	-0.0025 (0.0003)	-0.0025 (0.0003)
veteran effect on slope, β_1		-0.0035 (0.0023)	
veteran loss of experience, ℓ	2.08 (0.38)	1.84 (0.43)	
$\pi_1 = -[\beta_0\ell - \gamma\ell^2 + \beta_1\ell]$			-0.189 (0.052)
$\pi_2 = -[2\gamma\ell - \beta_1]$			0.006 (0.004)
Veteran overtaking age implied by reduced form estimates			50.1 (15.9)
χ^2(dof)	1.41 (1)	-	813.57 (1247)

Notes: Sample includes FICA taxable earnings from 1975-84 for men born 1950, 1976-84 for men born 1951 and 1977-84 for men born 1952.
Estimation method is described in Appendix B.
Asymptotic standard errors in parentheses.

Appendix A: Data and Standard Errors for the Adjusted FICA Earnings Series

1. Data

Mean earnings above the FICA taxable maximum (μ_j^ℓ) are taken from the 1974-85 Mare-Winship March Current Population Survey Uniform files. Years of birth were determined on the basis of age in 1985; ages 34-35 were assigned birth year 1950, ages 33-34 were assigned 1951, ages 32-33 were assigned 1952 and ages 31-32 were assigned 1953. Thus, sample cohort means are correlated because of double counting of men aged 32, 33 and 34. Table A.1 contains sample cell statistics for the four cohorts combined. Because of double counting when averaging across cohorts, the figures in the table represent weighted statistics (men aged 32-34 are given a weight of two and all others a weight of one). Column (1) shows the Social Security taxable maximum, column (2) shows the CPS censoring point, columns (3) and (4) show the weighted CPS mean and standard deviation of earnings above the Social Security taxable maximum, column (5) shows the weighted number of men above the Social Security maximum and column (6) shows the weighted total number of men in the CPS born in 1950-53.

2. Standard Errors

The censored FICA earnings are adjusted using

$$\mu_j^0 = \mu_j^c + p_j^{x\ell}(\mu_j^\ell - L_j).$$

Let $m = [m_1{}' \; m_2{}' \; m_3{}']'$ denote the vector of sample moments corresponding to μ^c, $p^{x\ell}$ and μ^ℓ and let Σ_{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

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

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 derived below. A simplified procedure, described at the end of Appendix B, is used to estimated the off-diagonal elements (correlations) of Ω directly from the adjusted cell means. The simplified procedure is based on the assumption that the correlation structure does not vary by lottery number for a given cohort.

Define a Bernoulli random variable, h_i, that equals one if individual i had a single source of earnings above the FICA taxable maximum and takes on the value zero otherwise. For men in cell j, let y_i be uncensored earnings and let y_i^* be CWHS censored earnings. That is,

$$y_i^* = \begin{cases} L_j & \text{if } h_i = 1 \\ y_i & \text{otherwise.} \end{cases}$$

Let $m_j = [m_{1j}\ m_{2j}\ m_{3j}]'$ be the vector of sample moments for cell j, containing information on n_j individuals in the CWHS. Then

$$m_{1j} = \frac{\sum_{i\epsilon j} y_i^*}{n_j} \quad \text{and} \quad m_{2j} = \frac{\sum_{i\epsilon j} h_i}{n_j}$$

Assuming i.i.d. sampling within cells, the covariance of m_{1j} and m_{2j} is given by

$$\sigma_{12j} = (1/n_j)[E(h_i y_i^*) - E(h_i)E(y_i^*)],$$

where σ_{12j} is the diagonal element of Σ_{12} corresponding to the jth cell.
But $h_i = 0$ unless $y_i^* = L$, so

$$\sigma_{12j} = [p_j^{x\ell}(L - \mu_j^c)]/n_j.$$

Table A.1: Men Born 1950-53, March CPS Mean Earnings Above SSA Taxable Maximum

Race	Year	(1) SSA Maximum	(2) CPS Maximum	(3) Mean(μ^{ℓ})	(4) Std. Dev.	(5) # > Max	(6) Cell Total
White	73	10,800	50,000	13,227	4,149	523	7,360
	74	13,200	50,000	15,926	2,844	317	6,993
	75	14,100	50,000	17,580	5,090	546	7,232
	76	15,300	50,000	19,335	5,394	912	9,228
	77	16,500	50,000	20,422	4,574	1,225	9,096
	78	17,700	50,000	22,205	5,184	1,602	8,910
	79	22,900	50,000	28,428	5,961	1,226	10,373
	80	25,900	50,000	31,817	6,099	1,049	10,531
	81	29,700	75,000	36,842	9,524	850	9,122
	82	32,400	75,000	42,570	10,872	822	9,139
	83	35,700	75,000	46,963	10,917	743	9,050
	84	37,800	99,999	49,057	13,832	860	9,006
Nonwhite	73	10,800	50,000	12,720	1,721	25	904
	74	13,200	50,000	16,054	1,691	13	922
	75	14,100	50,000	15,889	958	25	933
	76	15,300	50,000	20,437	7,499	52	1,187
	77	16,500	50,000	20,976	5,479	64	1,080
	78	17,700	50,000	22,768	5,055	119	1,082
	79	22,900	50,000	27,697	6,585	69	1,329
	80	25,900	50,000	32,233	7,684	86	1,298
	81	29,700	75,000	37,969	9,784	73	1,227
	82	32,400	75,000	44,377	11,636	67	1,243
	83	35,700	75,000	42,199	3,339	36	1,094
	84	37,800	99,999	45,510	6,278	46	1,110

Source: 1974-1985 Mare-Winship March Current Population Survey Uniform Files. Statistics and counts are weighted - see Appendix A.1, above.

Appendix B: Two-Sample Instrumental Variables and the Draft Lottery

GLS on equation (3) is the same as Instrumental Variables using dummy variable instruments. Estimation of (3) is a two-sample problem because $\bar{y}_{ctj}$ and $\hat{p}_{cj}$ are taken from different samples. Let $\bar{y}$ denote the vector of $\bar{y}_{ctj}$ and let $\hat{p}$ denote the vector of $\hat{p}_{cj}$. These moments correspond to $X_1'y_1/n_1$ and $X_2'Z_2/n_2$ in the notation of Angrist (1989c). Assuming moments from the two samples are independent, results in Angrist (1989c) imply that the optimal two-sample IV minimand is

$$\Phi = V(\bar{y}) + \alpha^2 V(\hat{p}),$$

where V(.) denotes the appropriate covariance matrix.

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 correlation only between elements of the time series of earnings for a particular cohort. The intertemporal correlation structure is assumed constant across lottery number cells 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 set. This procedure is also applied to the adjusted FICA series described in the previous 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 1953. 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.

The quadratic earnings models described in section V are estimated by modifying the above procedure to allow the treatment effects multiplying $\hat{p}$ to be time varying. The reduced form treatment effects associated with the quadratic models are of the form $(\pi_1 + \pi_2 x_c)$. Let Π_c denote the vector of $\pi_1 + \pi_2 x_c$ 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 cell j of cohort c:

$$(1/n_j)\Pi_c\{e_T e_T'[\hat{p}_{cj}(1-\hat{p}_{cj})]\}\Pi_c',$$

where e_T is a vector of T 1's. In practice, Π_c is replaced by the weighted least squares estimates (weights are the inverse sampling variances of $\bar{y}_{ctj}$) of the reduced form equation, (9).

For purposes of computation, all variables used in the estimation were transformed by the square root of the appropriate estimated block of Φ^{-1}. Φ was tabulated using the SAS MATRIX Procedure and the optimal estimates were computed by unweighted nonlinear least squares using the SAS (Version 5.0) procedure SYSNLIN.

Appendix C: The Matched Continuous Work History Sample[27]

1. CWHS data collection and processing

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. The sample is stratified using the regional information coded in the Social Security 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, rather they file IRS form W-2 with the SSA on an annual basis. The SSA forwards the employer-filed W-2 forms to the IRS. The IRS, however, has continued to collect completed W-2 forms directly from the taxpayer.

The importance of annual reporting is that all earnings, including those above the FICA taxable maximum, are now reported to the SSA on form W-2. Furthermore, all employers are required to file W-2's 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. Apparently, the SSA vigorously enforces only those reporting requirements essential to the operation of the Social Security program. A further shortcoming of the W-2 series is the poor quality of the data during the

[27] The material in this section draws heavily on Buckler and Smith (1984), U.S. Department of Health and Human Services (1987) and personal correspondence and conversations with Warren Buckler and Cresston Smith.

years of transition to annual reporting. As noted in the text, the SSA's own actuaries continued to rely on statistics compiled from IRS taxpayer-filed W-2's until 1985.

2. Coverage and truncation of the earnings series.

FICA covered earnings include 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 Meyer (1985, Table 2.1) and reproduced below:

Employment Category	Total Workers	Covered Workers	Fraction Covered
Industry and commerce	77.2	77.0	99.7
Nonprofit Organizations	6.8	5.3	78.8
State, local government	14.8	10.1	68.2
Civilian federal Government	3.5	0.5	15.4
Military federal government	2.1	2.1	100.0
Farm	2.2	1.6	74.6
Domestic	2.0	0.6	28.1
Railroad	0.5	0.5	100.1
Self-employed	9.0	6.8	75.5
Total	118.2	104.6	88.5

The figures for state, local and civilian federal government employees imply that by 1981 roughly 58% of all civilian government workers were covered.

FICA taxable maximums appear in text Table 1. The combined effects of limited coverage and censoring at the taxable maximum are conveniently summarized by the percentage of all earnings which are reported to the SSA. These statistics, taken from Dept. of Health and Human Services (1987, Table 30), are:

Type of earnings	Year of earnings 74	75	76	77	78	79	80	81	82	83	84
Wage and Salary	87.2	86.6	86.4	86.8	86.2	89.2	89.7	90.4	90.5	91.1	91.4
Self-employment	65.0	61.9	63.2	65.7	63.8	69.3	73.3	74.9	76.3	77.3	79.5

The CWHS W-2 earnings series excludes earnings from self-employment. Unpublished estimates indicate that roughly 6.4 percent of men born between 1944-1953 with non-zero earnings in 1984 had self-employment earnings only. Other differences between FICA and W-2 earnings coverage are described in Millea and Kilss (1980).

3. The matched file for the Draft Lottery Project

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. At my request, 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 (1969-73) semiannual reports. A small number of individuals were discarded from the final data set because they had no usable information on either sex, race or exact date of birth.

The Tax Reform Act of 1976 essentially prohibits the release of data on individuals collected for tax purposes. Therefore, an aggregate data set was constructed from the matched CWHS micro data. The aggregate data set includes statistics for cells defined by race (white and nonwhite), year of birth, year of earnings, type of earnings (FICA or W-2) and 5 consecutive lottery numbers. The information released includes cell means, variances, counts, fraction at the FICA limit, fraction above the FICA limit, fraction zero and number deceased according to the Social Security Benefit Status Code.

Appendix D: The Matched Survey of Income and Program Participation[28]

1. SIPP data

The Survey of Income and Program Participation is a Census Bureau longitudinal survey of the civilian noninstitutional population. Data for the first wave of the first SIPP panel were collected from four rotation groups, each of which was interviewed in one of four consecutive months. During the SIPP interview, information is collected for the interview month as well as retrospectively for each of the four preceding months. Earnings information, however, is not collected for the interview month.

2. The matched file for the Draft Lottery Project

The SIPP public use tapes contain year and month of birth. A Census Bureau "in-house" version of the SIPP contains information on day of birth which is not released to the public. At my 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.

Wave I of SIPP Panel I, the only SIPP data set used here, contains information on earnings in late 1983 and early 1984.

3. SIPP variables

Vietnam Era veteran status is determined by the SIPP variables VETSTAT, which records veteran status and by U-SRVDTE which records the period of service.

[28] The material in this section draws on U.S. Department of Commerce (1985).

Unless otherwise indicated, the earnings variables used in Table 8 refer to the month immediately preceding the interview month for each rotation group. The variable used for monthly earnings is

WS1_2032, monthly earnings on the respondents main job.

Average hourly earnings were computed from WS1_2032 and from

WS1_2028, hourly wage if the main job is paid hourly, interview month

WS1_2024, hours worked per week on the main job

WS1_WKS4, weeks worked per month on the main job.

Average hourly earnings are estimated using the following formula:

If WS1_2028 is missing then average hourly earnings equals

WS1_2032/(WS1_2024*WS1_WKS4),

otherwise average hourly earnings equals WS1_2028.

References

Angrist, Joshua D. (1988), "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 Earnings," forthcoming in <u>Research in Labor Economics</u> Volume X, Edited by Ron Ehrenberg, Greenwich: JAI Press, Inc., 1989.

__________ (1989b), "Selection for Military Service in the Vietnam Era," Industrial Relations Section Working Paper No. 250, Princeton University, March 1989.

__________ (1989c), "Two-Sample Two-Stage Least Squares," Industrial Relations Section, Princeton university, Mimeo, March 1989.

__________ and Alan Krueger (1989), "Why Do World War II Veterans Earn More?," Industrial relations Section, Princeton University, Mimeo, February 1989.

Ashenfelter, Orley C. and David Card (1985), "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs on Earnings," <u>The Review of Economics and Statistics</u> 67: 648-660.

Baskir, Lawrence M. and William A. Strauss (1978), <u>Chance and Circumstance: The Draft, The War and the Vietnam Generation</u>, New York: Alfred A Knopf, 1978.

Berger Mark C. and B.T. Hirsch (1983), "The Civilian Earnings Experience of Vietnam-Era Veterans," <u>The Journal of Human Resources</u> 18(4) (Fall 1983): 455:479

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

Bureau of National Affairs (1988), <u>The Daily Labor Report</u> 10-5-88, 1988 No. 193: B-1 to B-7.

Card, Josefina J. (1983), <u>Lives After Vietnam: The Personal Impact of Military Service</u>, Lexington: D.C Heath and Company, 1983.

Card, David and Daniel Sullivan (1988), "Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment," <u>Econometrica</u> 56: 497-530.

Chamberlain, Gary (1982), "Multivariate Regression Models for Panel data," <u>Journal of Econometrics</u> 18: 5-46.

Cohany, Sharon R. (1988), "BLS Reports on Labor Market Situation Among Disabled Veterans of the Vietnam Era," <u>News</u>, Bureau of Labor Statistics, U.S. Department of Labor, USDL 88-489, Washington DC, September 30, 1988.

Cooper, Richard V.L. (1977), <u>Military Manpower and the All-Volunteer Force</u>, RAND Corporation, Report R-1450-ARPA, Santa Monica: 1977.

Council of Economic Advisors (1988), Economic Report of the President, February 1988, Washington DC: U.S. Government Printing Office, 1988.

Crane, Jon R. and David A. Wise (1987), "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 N. (1982), "Veteran Status as a Screening Device," American Economic Review 72 (March 1982): 133-42.

Durbin, J. (1954), "Errors in Variables," Review of the International Statistical Institute 22: 23-32.

Fienberg, Stephen E. (1971), "Randomization and Social Affairs: The 1970 Draft Lottery," Science 171 (January 22, 1971): 255-261.

Hearst Norman, Tom B. Newman and Stephen B. Hulley (1986), "Delayed Effects of the Military Draft on Mortality: A Randomized Natural Experiment," New England Journal of Medicine 314 (March 6, 1986): 620-4.

Heckman, James J. and Richard Robb (1985), "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.

Heckman, James J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System," Econometrica 46 (July 1978): 931-959.

House Armed Services Committee (1970), Special Subcommittee on the Draft, 91st Congress, 2nd Session, Hearings for a Review of the Administration and Operation of the Draft Law, Statement of R.G. Kelley, Assistant Secretary of Defence for Manpower and Reserve Affairs, p.12620-12724.

Kelejian, H.H. (1971), "Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables," Journal of the American Statistical Association 66 (June 1971): 373-384.

Meyer, Robert J. (1985), Social Security, Third Edition, Homewood, Illinois: Richard D. Irwin, Inc., 1985.

Millea, Mary T. and Beth Kilss (1980), "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 (1974), Schooling, Experience and Earnings, New York: National Bureau of Economic Research, 1974.

New York Times (1987), "Official Defends Volunteer Army Against a Possible Plea for Draft," by R. Halloran, March 25, 1987.

Newsweek (1988), "The Political Specter of Vietnam," August 29, 1988.

Newey, Whitney K. (1985), "Generalized Method of Moments Specification Testing," Journal of Econometrics 29 (1985): 229-256.

Rosen, Sherwin and Paul Taubman (1982), "Changes in Life Cycle Earnings: What Do Social Security Data Show?," Journal of Human Resources 17(3): 321-38.

Schwartz, Saul (1986), "The Relative Earnings of Vietnam and Korean-Era Veterans," Industrial and Labor Relations Review 39: 564-572.

Selective Service System (1986), "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.

__________ (1971), "Semiannual Report of the Director of Selective Service for the Period July 1 to December 31 1970," Washington DC: U.S. Government Printing Office, 1971.

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

Smith, James P. and Finis Welch (1979), "The Effect of Cohort Size on Earnings: The Baby Boom Babies Financial Bust," Journal of Political Economy 87,5 Part 2 (October).

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

Taussig, Michael K. (1974), "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.

U.S. Department of Health and Human Services (1987), Social Security Bulletin, Annual Statistical Supplement, U.S. Department of Health and Human Services, Social Security Administration, 1987.

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

Veterans Administration (1981a), Research Monograph 15: Educational and Income Characteristics of Veterans, March 1979, Veterans Administration, Reports and Statistics Service, Washington DC: U.S. Government Printing Office, August 1981.

__________ (1981b), Data on Vietnam Era Veterans, Veterans Administration, Office of Reports and Statistics, Washington DC: U.S Government Printing Office, September 1981.

Wald, Abraham (1940). "The Fitting of Straight Lines if Both Variables are Subject to Error," Annals of Mathematical Statistics 11: 284-300.

Wall Street Journal (1988a), "Quayle Episode Raises Long-Dormant Issue: Were You in the War ?," August 25, 1988.

__________ (1988b), "Veterans Give Bush, Quayle Warm Support," By Ellen Hume, August 23, 1988.

White, Halbert (1982), "Instrumental Variables Regression with Independent Observations," Econometrica 50 (March 1982): 483-499.