NBER WORKING PAPER SERIES

USING SOCIAL SECURITY DATA ON MILITARY APPLICANTS TO ESTIMATE THE EFFECT OF VOLUNTARY MILITARY SERVICE ON EARNINGS

Joshua D. Angrist

Working Paper No. 5192

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1995

This research was supported by grant SBR-9122627 from the National Science Foundation. The Institute for Research on Poverty at the University of Wisconsin also provided financial support. Special thanks go to Warren Buckler, Russell Hudson, and the staff of the Office of Research and Statistics at the Social Security Administration's Baltimore Headquarters for providing and assisting with Social Security earnings data. Thanks also go to Mike Dove and the staff at the Defense Manpower Data Center for providing and assisting with data on military applicants and to Michael Sarel for outstanding research assistance. Finally, thanks go to seminar participants at Hebrew University, Brown, Columbia, and the MIT-Harvard Econometrics workshop for comments, and especially to Guido Imbens for some unusually helpful discussions. This revision was written while the author was visiting MIT. The author bears sole responsibility for the contents of this paper. The views expressed here are not meant to reflect the views of any US government agency or employee. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #5192
July 1995

USING SOCIAL SECURITY DATA ON MILITARY APPLICANTS TO ESTIMATE THE EFFECT OF VOLUNTARY MILITARY SERVICE ON EARNINGS

ABSTRACT

This study uses Social Security data on the earnings of military applicants to the all-volunteer forces to compare the earnings of Armed Forces veterans with the earnings of military applicants who did not enlist. Matching, regression, and Instrumental Variables (IV) estimates are presented. The matching and regression estimates control for most of the characteristics used by the military to select qualified applicants from the military applicant pool. The IV estimates exploit an error in the scoring of exams used by the military to screen applicants between 1976 and 1980. All the estimates suggest that soldiers who served in the early 1980s were paid considerably more than comparable civilians while in the military. Military service also appears to have led to a modest (less than 10 percent) increase in the civilian earnings of nonwhite veterans while actually reducing the civilian earnings of white veterans. Most of the positive effects of military service on civilian earnings appear to be attributable to improved employment prospects for veterans.

Joshua D. Angrist
Department of Economics
Hebrew University of Jerusalem
Mount Scopus
Jerusalem 91905
ISRAEL
and NBER

1. Introduction

From 1989 to 1992, the number of yearly enlistments in the Armed Forces by men and women without prior military service fell by 27 percent (Angrist 1993a.) This decline was partly achieved by a increases in test-score cutoffs and other entry standards (Defense Department 1992.) The social implications of these reductions depend in large part on whether voluntary military service has had a beneficial or negative impact on soldiers' labor market outcomes (Laurence 1992.) The labor market consequences of military service are of especial interest for minority recruits because blacks disproportionately apply to enter the military and have been over-represented in active-duty forces since the early 1970s (Angrist 1991, Defense Department 1988). The importance of the link between military service and the civilian labor market notwithstanding, empirical evidence on the effects of voluntary military service remains scarce.[1]

In spite of the lack of evidence on the effects of voluntary military service, there is considerable evidence that being *drafted* into the military can actually hurt future earnings. In Angrist (1989, 1990), I used the Vietnam-era draft lottery to show that military service in the Vietnam era reduced the civilian earnings of white veterans, and probably had no beneficial effect on nonwhite veterans. Elsewhere, Alan Krueger and I (Angrist and Krueger 1994) used a similar instrumental variables strategy to show that even though simple comparisons suggest that World War Two veterans earn more than nonveterans of the same

[1]Two recent studies of the effects of voluntary military service have been limited to small samples and a single year of earnings data from the National Longitudinal Survey of Youth (Bryant, Samaranayake, and Wilhite 1993, Magnum and Ball 1989).

age, the causal effect of military service in World War Two is probably negative.[2]

Although evidence from draft-era studies present a consistent picture of mostly negative effects of military service on earnings, results from these draft-era studies might not apply to veterans of the All-Volunteer Forces (AVF). Since 1973, new recruits to the US Armed Forces have freely chosen to enter the military, partly in response to generous packages of compensation and benefits. Some veterans benefits, such as subsidies for schooling, appear to improve the civilian labor market outcomes of veterans (Angrist 1993b). Of course, the fact that veterans benefits make veterans better off does not imply that veterans are better off than comparable civilians. Even though they are volunteers, AVF recruits could fail to benefit from their service if they do not accurately evaluate future costs and benefits at the time of enlistment.

This paper presents new evidence on the effects of voluntary military service on the earnings of veterans who applied to enter the military between 1979 and 1982. To carry out this study, I matched administrative data for a random sample of nearly 700,000 military applicants to their Social Security earnings histories from 1974 through 1991. The resulting data set contains demographic and other information collected at the time of application, as well as information on veteran status and earnings. Typically, only half of military applicants actually enter the military. Data on non-enlisting military applicants therefore

---

[2]Other studies of the effects of compulsory military service include Cutright (1974), who linked 1964 Social Security earnings to data on conscripts and a matched control group. Cutright's findings suggest a negative impact of military service for white veterans and no effect for nonwhite veterans. Laurence, Ramsberger, and Gribben (1989) found that low-ability volunteers from the late 1970s (a group examined here as well) and low-ability veterans from the Vietnam-era "Project 100,000" group did not benefit from military service. On the other hand, Beusse's (1974) earlier findings using Social Security data suggest that low-ability veterans from "Project 100,000" did earn more than comparable civilians. Studies of Vietnam veterans by Berger and Hirsch (1983) and DeTray (1982) report both positive and negative effects.

provide a natural control group for estimating the effects of military service. Although rejected or qualified but not-enlisting applicants are not a random sample of all applicants, these groups probably come closer than comparison samples from survey data to providing information on what the outcomes of military entrants would have been if they had not served in the military.

The paper begins with matching and regression estimates that fully control for all observed differences between enlisting and non-enlisting applicants. The matching estimates are generated from cell-by-cell comparisons of veteran and nonveteran earnings. These comparisons are then weighted to produce estimates of the effect of treatment on the treated as suggested by Rubin (1977). In addition, the paper presents IV estimates based on a strategy which exploits changes in the way military entrance examinations were scored between 1979 and 1982. When the Armed Forces Vocational Aptitude Battery (ASVAB) was first introduced as an applicant screening test in 1976, the score scales used to grade applicants unintentionally allowed large numbers of previously unqualified applicants to enter the military. This episode has come to be known as the "ASVAB misnorming." Correction of the ASVAB score scale in October 1980 reduced the probability of acceptance to the military by as much as 30 percent for some low-scoring military applicants. This abrupt increase in rejection rates is used here to construct instrumental variables that are highly correlated with veteran status in the applicant population.

The next section reviews aspects of the military record-keeping system and military administrative records. Section 3 describes the Social Security earnings data. Section 4 presents the matching and regression estimates of the effects military service. Section 5 presents the IV estimates and Section 6 concludes.

2. Data on military applicants

The military data analyzed here come from DMDC files containing information on applicants and entrants to the military for each fiscal year.[3] These records indicate the date and type of applicants' contacts with the military, and they include information collected at the time of contact. Possible types of contact include application and examination, enlistment into the military's Delayed Entry Program (DEP), enlistment to active duty, or discharge from DEP without active duty enlistment. Each record also reports basic demographic information, physical examination results, ASVAB test scores, and the terms and conditions of enlistment for new entrants to the military.

DMDC applicant records do not indicate whether an applicant eventually enlists. Instead, each enlistment generates a new record in the DMDC filing system. All records, however, contain Social Security Numbers (SSNs) that I have used to link information on applicants with information on entrants. Details of this link are described in my earlier paper on military applicants (Angrist 1993a.) Briefly, applicants for whom an entrants record was found in the year of application or in the following two years (and who were recorded on the entrants' record as entering active-duty service) were identified as veterans.

After matching data on applicants to data on entrants, I matched this linked data set to Social Security earnings histories. The sample matched to earnings was limited to male

---

[3]Aspects of the DMDC record-keeping system are described in Berryman, Bell, and Lisowski (1983) and Orvis and Gahart (1990).

applicants who applied during calendar years 1976-82, were aged 17-22 at the time of application, had valid sex and race codes, had data on Armed Forces Qualification Test (AFQT) scores collected on ASVAB test forms 5/6/7 or later,[4] and had at least a 9th grade education but no more than a 4-year college degree at the time of application. I also dropped observations on entrants who appeared to enter the military before their date of application, or who waited more than two years between application and entry to the military. Only the most recent record was retained in case of duplicates.

The target applicant population contains 2,161,000 white men and 862,576 nonwhite men. From this group, a random sample that includes roughly 15 percent of white applicants and 46 percent of nonwhite applicants was selected. The sample was designed to be slightly more than half nonwhite and 93 percent self-weighting. Certain groups were over-sampled so as to satisfy Social Security Administration (SSA) confidentiality requirements. Additional details on the sample design and Social Security confidentiality review are provided in the data appendix. The sample submitted to SSA to be matched to earnings includes 753,095 observations, of which 352,035 are white applicants and 401,060 are nonwhite applicants.

2.1 Descriptive Statistics for the population and sample

Table 1 provides descriptive statistics for the applicant population from which the sample linked to SSA data was drawn. Also shown is the proportion veteran by year of

---

[4]The AFQT is a widely-used composite test score that appears in the ASVAB test battery as a collection of subtests. The applicant population to be matched was restricted to those tested on ASVAB Form 5/6/7 or later so that AFQT scores would be roughly comparable across years, and because this form was the first used extensively in the ASVAB testing program (Maier and Truss 1983.)

application and race. The large increase in the number of applicants in 1980 is probably the result of increasingly generous veterans benefits and recruiting efforts associated with the then-incoming Reagan administration's defense build-up. The fraction of white applicants who ended up enlisting ranges from a high of 55 percent in 1979 to a low of 49 percent in 1981. The corresponding figures for nonwhite applicants range from a high of 50 percent in 1978 to a low of 36 percent in 1981. The 1981 decline in the fraction of nonwhite applicants who enlisted is the result of the increase in test score cutoffs that came about when the ASVAB score scale was corrected in 1980 (Eitelberg, *et al* 1984; Angrist 1993a.)

Table 2 shows the size and proportion veteran by year of application in the sample of 697,944 applicants for whom Social Security earnings records were located. The proportion veteran is slightly higher in the matched sample than in the population. This may be because veterans are more likely than nonveterans to have ever been in FICA-covered employment.

Statistics not shown in the tables indicate that roughly 30 percent of applicants in the sample were 18 when they applied to the military, 25 percent were aged 19, and 16 percent were aged 20. Only 13 percent were aged 17 and 16 percent were aged 21 or 22. A total of 40 percent of the applicants in the sample were high school graduates at the time of application. An additional 4 percent were GED certified and 34 percent had completed 11th grade. 18 percent had no schooling beyond 9th or 10th grade and only 4 percent had some college. Out of nearly 700,000 applicants in the sample, only 739 were college graduates.

Table 3 shows the distribution of AFQT scores in the applicant sample by year of application and race. AFQT percentile scores are derived from a score scale that indicates how well an individual did relative to a reference population. These percentile scores are determined in a process called test-equating or test-norming (Holland and Rubin 1982). AFQT scores are usually reported on a scale of I to V, where AFQT group I denotes percentiles 93-100, AFQT group II denotes percentiles 65-92, AFQT group IIIa denotes percentiles 50-64, AFQT group IIIb denotes percentiles 31-59, AFQT group IV denotes percentiles 10-30, divided into three subgroups (IVa, IVb, and IVc,) and AFQT group V denotes the lower decile in the reference population.

The AFQT score groups shown in Table 3 and throughout this paper are based on a common score scale for the whole sample period. This scale was derived from a reference group of World War II conscripts tested in 1944.[5] Because the AFQT scores reported in Table 3 are based on a consistent set of 1944 test norms, the same score should indicate roughly the same level of ability in any year in the sample.

The score distributions in Table 3 show that white applicants have higher scores than nonwhite applicants. In 1976, for example, 21 percent of nonwhite applicants were in AFQT group V, while only 6 percent of white applicants were in this group. Applicants with AFQT scores in group V are legally barred from enlisting. The score distribution also shows a significant increase in scores over time. By 1982, the proportion of applicants in score group V fell to almost one-third of its 1976 level. The proportion with scores above the 50th

---

[5]DMDC data on AFQT scores for test-takers in 1976-80 include up to 3 score scales. The first is the score scale in use at the time, which was later shown to be incorrect. A corrected score scale based on the old 1944 reference population was introduced in late 1980. Also in 1980, a new score scale was developed when a nationally representative sample of high school students was given the ASVAB test. Score scales based on these 1980 norms became standard in October 1984. For men applying after 1980, I converted the 1980 score scale to a corrected 1944 score scale using unpublished DMDC conversion tables. Thus, all the AFQT scores used here are based on the corrected 1944 score scale.

percentile (AFQT group IIIa and above) increased from 46 percent to 52 percent of white applicants and from 13 percent to 19 percent of nonwhite applicants. One reason for the increase in scores is an improvement in the package of veterans benefits and pay offered to new recruits in the early 1980s (Angrist 1993b.) In addition, military recruiters were increasingly likely to pre-screen applicants using a mini-AFQT exam. High school dropouts with low preliminary scores are discouraged from applying to the military, especially when recruiting conditions are good (Eitelberg, *et al,* 1984.)

3. Social Security earnings data

The SSA keeps track of the earnings and insured status of all FICA-covered workers in the United States in a data base called the Summary Earnings Record (SER). SSA programmers were able to locate earnings records in the SER for 697,944 applicants in the target sample of 753,095.[6] Observations on FICA earnings for each year from 1974-91 were then added to the applicant sample, generating a micro data set containing information on applicants, veteran status, and earnings.

For persons without earnings from FICA-covered employment in a given year, the SER shows zero earnings for that year. An individual who has *never* worked in FICA-covered employment probably has no record in the SER. But almost all non-governmental employees are covered by the main FICA programs, OASDI (old-age, survivors, and disability insurance) and HI (health insurance or Medicare). Members of the armed services have also been covered by FICA since 1956, and most federal employees have been covered

[6] 90 percent of the observations in the matched sample are self-weighting.

since 1983. State and local employees have been able to elect FICA coverage since 1954, and have been covered by Medicare since 1986 and optionally by OASDI since 1990. FICA earnings data on the SER for each year are censored at the FICA taxable maximum for that year. Additional information on coverage and censoring is given in the data appendix.

The SSA does not release individual earnings data to researchers. The micro data set linking 697,644 military applicants to their earnings was therefore used to produce a data set containing average earnings for each of 8,760 cells defined by race, year of application, AFQT score group, veteran status, schooling level at the time of application, year of birth, and a qualification-status variable for applicants in 1977-78 which was not used in this project. A description of these cell-identifiers appears in the data appendix.

Of the 8,760 cells in the population, 5,654 had 25 or more observations. The released data set contains cell-identifiers for these 5,654 cells, along with average FICA earnings for 1974-91. Additional earnings variables provided for each cell include: average positive earnings, average log earnings, the corresponding standard deviations, the fraction in the cell with zero earnings, the fraction in the cell with earnings exactly at the FICA taxable maximum, and the fraction of the cell with earnings above the FICA taxable maximum.[7] The data released by the SSA were subject to a confidentiality-edit which masks some or all of the earnings variables in small cells. A detailed description of the confidentiality edit is also provided in the data appendix.

---

[7]Persons with multiple employers can have earnings above the FICA taxable maximum on the SER.

3.1 Earnings profiles

Figures 1a and 1b show average earnings profiles by application-year cohort for military applicants who applied in 1976-82 (earnings are in 1991 dollars.) The points plotted in these figures are weighted averages over cells, using the cell sample count as weights. Figure 1a plots earnings profiles for whites and Figure 1b plots earnings profiles for nonwhites. The figures show a high rate of initial earnings growth, with declining and even negative growth rates later on. The profiles also show an increase in the rate of earnings growth in the year of application, probably attributable to a reduction in the number of men with zero earnings in that year. The dip in earnings in 1980-82 is a business cycle effect that appears to have hit older and nonwhite workers the hardest.

Figures 2a and 2b reproduce the earnings profiles for positive earnings only. These figures show a more modest slowdown in earnings growth during the 1980-82 recession and lower rates of earnings growth in other years. Even in the sample of positive earners, the profiles exhibit a concavity familiar from studies of labor earnings over the life-cycle. In this case, however, at least part of the decline in earnings growth at the end of sample period is attributable to the lags with which data on Social Security earnings are recorded in the SER data base.[8]

In the remainder of this paper, I focus on 1979-82 applicants with AFQT scores in groups III and IV (excluding a handful of college graduates.) The reasons for this focus are, first, that the applicant group most affected by recent increases in military entrance standards

[8]The problem of late reporting of Social Security earnings is discussed in the appendix to Angrist (1990) and by Card and Sullivan (1988).

consists of those with AFQT scores in group IV and, to a lesser extent, group III.[9] Second, in addition to being of special policy interest, men with AFQT scores in groups III and IV constitute the majority of applicants. For example, men scoring in AFQT categories III and IV constitute 64 percent of sampled white applicants in 1982 and 84 percent of sampled nonwhite applicants in 1982. Finally, applicants from 1979 onwards are of particular interest because, beginning in 1979, applicants were offered increasingly generous packages of veterans benefits that remain in place today.[10]

Figures 3a and 3b plot average earnings profiles by veteran status for the sample of 1979-82 applicants with AFQT scores in groups III and IV. The averages were constructed from cell means using population cell counts as weights. Table 4 reports the difference in earnings by veteran status for the sample underlying Figures 3a and 3b, along with standard errors. Panel A of Figure 4 is for all earnings and Panel B is for positive earnings. Column 2 in Table 4 shows the veteran earnings gap for whites and column 5 shows the veteran earnings gap for nonwhites. For comparison, columns 1 and 4 report mean earnings. Standard errors for the differences are shown in columns 3 and 6.

Differences in earnings by veteran status before 1979 are small but, in some cases,

---

[9]In 1989, there were 17,700 enlistments by men with AFQT scores in group IV but in 1992 there were only 340 enlistments by men with AFQT scores in group IV. Similarly, enlistments by applicants scoring in AFQT group IIIb fell by almost 40 percent over the same period (Angrist 1993a, Table 2.)

[10]The main benefit for schooling between 1976 and 1982 was offered through the Veterans Educational Assistance Program (VEAP.) The VEAP offered benefit levels considerably lower than the benefits offered under the previous Post-Korean GI Bill. Beginning in 1979, however, the individual services (especially the Army) began to offer a range of additional benefits which, in combination with VEAP, approached the benefit levels in previous and later programs (Angrist 1993b.)

statistically significant. After 1977, veterans earn more than nonveterans. The gap in favor of white veterans reaches over 1500 dollars in 1982 and 1991, which is over 10 percent of average FICA earnings for whites at the end of the sample period. The largest veteran premium for nonwhites is roughly 2900 dollars in 1982 and 1983. The next largest estimate for nonwhites is almost 2800 dollars in 1991, which amounts to over 24 percent of average FICA earnings for nonwhites at the end of the sample period. Most veterans of the AVF served for roughly 3 years (though many serve longer), so that the differences in earnings by veteran status between 1980 and 1985 largely reflect differences between military and civilian earnings in those years.[11]

Panel B of Table 4 shows estimates of the veteran earnings gap for men with positive earnings. The contrast between Panels A and B suggests that a large part of the veteran earnings advantage is an employment effect or a FICA-coverage effect. But conditional on having positive earnings, the veteran premium remains large in some years, especially for nonwhites. The largest positive earnings gap for whites is over 700 dollars in 1982 and almost 900 dollars in 1991. The largest gaps for nonwhites are between 1800-1900 dollars in 1982, 1983, and 1991.

---

[11]My tabulations from the 1987 Survey of Veterans show that nearly 60 percent of AVF veterans in 1987 served less than 5 years on active duty. Most veterans in the sample of 1979-82 applicants were therefore discharged in the mid 1980s. Over 95 percent of AVF veterans served less than 10 years.

4. Matching and regression estimates of the veteran earnings gap

4.1 Matching estimates

The primary purpose of this paper is to estimate the impact of military service on the earnings of applicants for military service. In particular, attention is focused on the difference between the earnings of veterans and what these veterans would have earned if *they* had not served in the military. In the program evaluation literature, this difference is called the effect of treatment on the treated (Heckman and Robb, 1985) or the selected average treatment effect (Angrist and Imbens, 1991).

The parameter of interest can be formally described as follows. For any applicant observed after application, define random variables representing what the applicant would earn had he served in the military and what the applicant would earn had he not served in the military. Denote these two potential outcomes by Y_0 and Y_1 and denote veteran status by a dummy variable, D. For each applicant, we observe only $Y = Y_0 + (Y_1-Y_0)D$, although it is Y_1-Y_0 that we are interested in. Because the difference Y_1-Y_0 is never observed for any individual, the object of estimation is an average of Y_1-Y_0. The effect of treatment on the treated is one such average,

$$E[Y_1 - Y_0 | D=1] = E[Y_1 | D=1] - E[Y_0 | D=1].$$

This tells us whether, on average, a given group of veterans benefitted or suffered from military service.[12]

The comparisons in Table 4 provide one estimate of $E[Y_1 - Y_0 | D=1]$. Because the

[12]The formulation of evaluation problems in terms of averages of potential outcomes is usually attributed to Rubin (1974, 1977).

underlying sample includes only applicants to the military, these comparisons control for differences between veterans and nonveterans that originate in the decision to apply to the military. But the comparisons in Table 4 do not control for most of the criteria used by the military to choose which applicants to accept. The earnings of non-enlisting applicants might therefore provide a poor indicator of what enlisting applicants would have earned if they had not served in the military. To explore this point further, note that the comparisons in Table 4 provide estimates of the difference between $E[Y_1| D=1]$ and $E[Y_0| D=0]$ in the applicant pool. This difference can be decomposed as follows:

(1) $E[Y_1| D=1] - E[Y_0| D=0] = E[Y_1 - Y_0| D=1] + \{E[Y_0| D=1] - E[Y_0| D=0]\}$.

Equation (1) shows that comparisons of earnings by veteran status are equal to $E[Y_1 - Y_0| D=1]$ plus a bias term. This bias term is attributable to the fact that even in the applicant population, the earnings of nonveterans are not necessarily representative of what veterans would have earned had they not served in the military.

If veteran status was randomly assigned in a social experiment, then D would be independent of Y_0, implying $E[Y_0| D=0] = E[Y_0| D=1]$. Rosenbaum and Rubin (1983) refer to treatment status that is independent of potential outcomes as an *ignorable* treatment assignment. Although claims for ignorable treatment assignment are usually implausible outside of an experimental setting, it is plausible that veteran status among applicants is ignorable conditional on a set of observed covariates. In particular, it might be the case that in the pool of military applicants the principle sources of bias in veteran/nonveteran comparisons are differences in observed variables such as age, schooling, and test-scores. These are the main variables used by the military to screen and select applicants and they are observed in the DMDC/SSA matched data set.

The assumption that veteran status is ignorable conditional on covariates, denoted by X, can be expressed formally as follows:

(2) $E[Y_0 | X, D=1] = E[Y_0 | X, D=0]$.

In other words, conditional on X, observations on nonveterans provide a valid control group for veterans.[13] When (2) holds, the effect of treatment on the treated can be estimated using the sample analog of the following equation:

$$(3) \quad E[Y_1 - Y_0 | D=1] = E\{E[Y_1 | X, D=1] - E[Y_0 | X, D=0] | D=1\}$$
$$= \int \{E[Y_1 | X, D=1] - E[Y_0 | X, D=0]\} P(X | D=1) dx$$

This expression means that $E[Y_1 - Y_0 | D=1]$ is obtainable from a weighted average of contrasts between veteran and nonveteran earnings at each value of X. The weights are given by the density or distribution function of X among veterans. Values of X for which there are no veterans, i.e., for which $P(D=1 | X)=0$, are given zero weight.

A problem with this approach arises when there are values of X at which $P(D=1 | X)=1$. At such values, $E[Y_0 | X, D=0]$ is undefined. This problem is resolved here by (implicitly) defining the parameter of interest to be $E[Y_1 - Y_0 | D=1, P(D=1|X)<1]$. In some cases, this additional restriction is not binding. For example, if $P(D=1 | X)$ is generated by a latent index model, then we have $0 < P(D=1 | X) < 1$ with strict inequalities for all values of X. In other cases, imposing $P(D=1 | X) < 1$ restricts the

[13]Equation (2) has been called the "selection on observables" assumption in a regression context (see, e.g, Barnow, Cain, and Goldberger, 1981). In this case, however, neither linearity or constant treatment effects assumptions are imposed. Regression-control for observables differs from the non-parametric control strategy discussed here. See Dehejia and Wahba (1995) for a recent implementation of the nonparametric approach in a manpower training evaluation.

parameter of interest to a population where groups of veterans and nonveterans are available for comparison at each value of X.

To construct an estimator based on (3), suppose that X is discrete, taking on values $\{x_1, \ldots, x_k, \ldots, x_K\}$. Suppose also that there are N_{1k} observations in the population of veterans with $X=x_k$ and N_{0k} observations in the population of nonveterans with $X=x_k$. The corresponding sample sizes are denoted n_{1k} and n_{0k}. Let $\delta_k = 1[n_{1k}>0, n_{0k}>0]$. Finally, let $\bar{y}_{1k}$ denote the average earnings of veterans with $X=x_k$ and let $\bar{y}_{0k}$ denote the average earnings of nonveterans with $X=x_k$. Then the following estimator is unbiased for $E[Y_1 - Y_0| D=1, P(D=1|X)<1]$ when equation (2) holds (Rubin, 1977):

$$(4) \qquad \hat{\alpha}_c \equiv \Sigma_k\ \delta_k N_{1k}[\bar{y}_{1k} - \bar{y}_{0k}]/\Sigma_k\ \delta_k N_{1k};$$

This estimator shares with (3) the property that earnings comparisons are made only at values of X where both veterans and nonveterans are observed.[14] Comparisons at each value of X are then weighted by $\delta_k N_{1k}/\Sigma_k\ \delta_k N_{1k}$, which is the population distribution function of X among veterans at values of X where $[\bar{y}_{1k} - \bar{y}_{0k}]$ is defined. In what follows, I refer to $\hat{\alpha}_c$ as a controlled contrast because it is an estimator that controls for all observed differences between veterans and nonveterans. In practice, the observed covariates take on values in the set of all possible combinations of race, application year, schooling at the time of application, AFQT score group, and year of birth.

Figures 4a and 4b plot estimates of $\hat{\alpha}_c$ for the earnings of white and nonwhite

[14]The estimator $\hat{\alpha}_c$ differs slightly from the one discussed by Rubin (1977) in that it allows for non-self-weighting samples. Since population cell sizes are known here they can be used as weights. More generally, comparisons of veteran and nonveteran earnings could be weighted by n_{1k}/f_k, where f_k is the sampling rate at $X=x_k$.

applicants from 1974-91. Also shown in the figure is a simple comparison of veteran and nonveteran earnings for the same sample underlying estimates based on (3). These comparisons are estimated as follows:

(4) $\hat{\alpha}_n \equiv \{\Sigma_k \delta_k N_{1k}\bar{y}_{1k}/\Sigma_k \delta_k N_{1k}\} - \{\Sigma_k \delta_k N_{0k}\bar{y}_{0k}/\Sigma_k \delta_k N_{0k}\}$.

Because $\hat{\alpha}_n$ does not control for observed covariates, I refer to it as a naive contrast. Note that the only difference between $\hat{\alpha}_n$ and the estimated veteran/nonveteran earnings differences reported in Table 4 is that $\hat{\alpha}_n$ excludes cells where either $n_{1k}=0$ or $n_{0k}=0$ in the sample.

Figures 4a and 4b plot estimates of $\hat{\alpha}_c$ and $\hat{\alpha}_n$ along with 95 percent confidence bands. Estimates of $\hat{\alpha}_c$ and the corresponding standard errors are also reported in Table 5. The estimation of standard errors is discussed in the statistical appendix. The estimates of $\hat{\alpha}_c$ (labelled "controlled contrast" in the figure) and $\hat{\alpha}_n$ (labelled "naive contrast" in the figure) are positive for both racial groups from 1980-84, but $\hat{\alpha}_c$ is much less than $\hat{\alpha}_n$ in every year. The largest estimate of $\hat{\alpha}_c$ is 783 dollars for whites and 2186 dollars for nonwhites, both in 1982. These estimates amount to just over 10 percent of white applicants' average 1982 FICA earnings and over 35 percent of nonwhite applicants' average 1982 FICA earnings. The associated standard errors are on the order of 35 dollars. The graph of $\hat{\alpha}_c$ shows only small (though sometimes statistically significant) differences in earnings by veteran status for white applicants until 1980 and for nonwhite applicants until 1979. The lack of substantial pre-application differences in earnings is important evidence in favor of a causal interpretation of the controlled contrast.

Beginning in 1983 the estimates of $\hat{\alpha}_c$ fall, becoming negative for whites in 1984 and remaining negative through 1990. The 1991 estimate for whites is about 30 dollars but this amount is not statistically different from zero. Estimates for nonwhites remain positive throughout the sample period but fall to less than 1000 dollars from 1985 through 1990. The 1991 estimate of 1026 dollars for nonwhites is less than 9 percent of nonwhite's average 1991 FICA earnings.

One important implication of Figures 4a and 4b is that the naive contrast is a highly misleading over-estimate of the benefits of military service. This is especially true after 1985, by which time most veterans have completed their military service. For example, for nonwhites in 1991, $\hat{\alpha}_n$ is almost triple $\hat{\alpha}_c$. Similarly, although $\hat{\alpha}_n$ is never negative for white applicants, $\hat{\alpha}_c$ is negative for whites in every year after 1983 except 1991. The gap between $\hat{\alpha}_n$ and $\hat{\alpha}_c$ is largely explained by the fact that among applicants, veterans have higher test scores and are more likely than nonveterans to be high school graduates.

4.2 Additional matching estimates

Figures 5a and 5b plot estimates of $\hat{\alpha}_c$ by race and application year for single application-year cohorts applying in 1979-82. Differences in the time-series pattern of $\hat{\alpha}_c$ by application year suggest that nonzero estimates of $\hat{\alpha}_c$ are attributable to military service. In particular, for both race groups, $\hat{\alpha}_c$ is close to zero until the year of application, at which time there is a sharp break. After this break, which is especially well-defined for nonwhite applicants, veterans earn considerably more than nonveterans for a few years. Similarly, the decline in $\hat{\alpha}_c$ after 1982 is spaced at one-year intervals across application-year cohorts. The timing of these declines suggest that they are generated by a staggered return to civilian life by veterans in each application-year cohort.

Figure 6 and Panel B in Table 5 report estimates of $\hat{\alpha}_c$ for positive earnings.[15] Estimates for whites and nonwhites are plotted on the same axis. The estimates of $\hat{\alpha}_c$ in Figure 6 were computed with N_{1k} equal to the sample count of veterans with positive earnings at $X=x_k$, inflated by the inverse sampling rate at that value of X. As with the naive contrast in Table 4, Figure 6 suggests that much of the positive impact of veteran status in controlled contrasts is attributable to improved employment prospects for veterans. Conditional on working, $\hat{\alpha}_c$ for whites is negative in every year except 1982, when it is only 102 dollars. After 1983, $\hat{\alpha}_c$ for whites ranges from -582 dollars to -1166 dollars. Estimates of $\hat{\alpha}_c$ for nonwhites for 1986 through 1990 are also negative, ranging from -100 to -175 dollars. The 1991 estimate for nonwhites is only 70 dollars, with a standard error about the same size.

Figure 7 plots estimates of a controlled contrast for average log earnings. The sample underlying these estimates is the same as that underlying Figure 6 (i.e., men with positive earnings) but the estimates show a slightly different pattern than those in Figure 6. Although $\hat{\alpha}_c$ is large and negative for the positive earnings of whites after 1984, the corresponding estimates of $\hat{\alpha}_c$ for log earnings are small and positive or zero. Similarly, while $\hat{\alpha}_c$ is negative or close to zero for the positive earnings of nonwhites after 1984, the corresponding estimates of $\hat{\alpha}_c$ for log earnings are .10-.15. In other words, estimates of $\hat{\alpha}_c$ for log earnings suggest a 10-15 percent earnings premium for nonwhite veterans even though the difference in average positive earnings by veteran status is zero.

[15]Because employment status is also an outcome variable, estimates of treatment effects for earnings conditional on employment status do not necessarily have a causal interpretation, even when the unconditional estimates do (see, e.g., Angrist 1995).

The contradiction in results for positive earnings and log earnings is explained by Figure 8. Points plotted in Figure 8 were computed by replacing mean earnings with the variance of earnings in the formula for $\hat{\alpha}_c$. This figure shows that veteran status is associated with a large and persistent reduction in the standard deviation of earnings. The importance of Figure 8 for Figure 7 is that expected log earnings is a declining function of the variance of earnings. In particular,

$$E[\ln y] \approx \ln[E(y)] - 1/2\sigma^2/E(y)^2,$$

where σ^2 is the variance of earnings. Reductions in the variance of veterans' earnings therefore serve to raise the mean of veterans' log earnings. The reason veteran earnings have lower variance is probably because there is little dispersion in earnings for men on active duty who began military service at same time. On the other hand, the reduction in variance persists beyond the years when most veterans were in the military.

4.3 Regression estimates

A natural question raised by the approach in Sections 4.1 and 4.2 is whether a more conventional (in econometrics) regression strategy generates similar estimates. As before, let X denote a discrete covariate that indicates all unique combinations of year of birth, schooling, application year, and AFQT score group. Regression estimates of the effect of military service are based on the following model for each year and race group:

$$\text{(5)} \qquad \bar{y}_{jt} = m(X_j)\beta + \alpha_r D_j + \bar{\epsilon}_{jt}$$

where $\bar{y}_{jt}$ is cell earnings, $m(X_j)$ is the jth row of the design matrix for a saturated model for X with coefficient vector β, α_r is the regression veteran effect, and ϵ_j is an error term that is orthogonal to $m(X_j)$ and D_j by construction. The fact that $m(X_j)$ is saturated ensures that there will be no omitted variables bias arising from any association between D_j and X_j.

The results of estimating equation (5) by weighted least squares are plotted in Figure 9. The weights are the population cell counts (N_{1k} for veterans, N_{0k} for nonveterans) to correct for nonrandom sampling; the regression itself corrects for differences in the mean of X between veterans and nonveterans. The regression estimates are therefore based on a sample that includes all cells with any observations on veterans or nonveterans. For comparison, the regression estimates are overlaid with the corresponding set of matching estimates. The matching estimates plotted in Figure 9 are the same as those plotted in Figures 4a and 4b.

The regression and matching estimates are virtually identical for each year from 1974 through 1984. But the regression estimates for each year after 1984 are larger than the matching estimates, especially for nonwhites. For the 1991 earnings of nonwhites, the regression estimates exceed the matching estimates by almost 25 percent.

Why do the regression estimates tend to over-estimate the effect of treatment on the treated in this application? This question can be answered with the help of a simple example involving one binary covariate, X_1. Subscripts and overbars are dropped for this illustration. In this example, the matching and regression estimates can both be described in terms of the parameters of the following regression:

(6) $$Y = \beta_0 + \beta_1 X_1 + \alpha_0 D(1-X_1) + \alpha_1 DX_1 + \epsilon.$$

Because this model is saturated in both X_1 *and* D, it can be used to construct estimates of $E[Y_1 - Y_0 | D=1]$ when assumption (2) holds. In particular,

$$E[Y_1| X_1=1, D=1] - E[Y_0| X_1=1, D=1] = \alpha_1,$$

and

$$E[Y_1| X_1=0, D=1] - E[Y_0| X_1=0, D=1] = \alpha_0,$$

so that,

$$E[Y_1 - Y_0| D=1]$$

$$= \alpha_0 P[X_1=0| D=1] + \alpha_1 P[X_1=1| D=1]$$

$$= \{\alpha_0 P[D=1| X_1=0]P[X_1=0] + \alpha_1 P[D=1| X_1=1]P[X_1=1]\}/P[D=1]. \tag{7}$$

Unlike equation (6), the model used to construct the regression estimates plotted in Figure 9 ignores interactions between X and D. Similarly, suppose the underlying conditional expectation function is given by equation (6), but we ignore interaction terms and instead estimate

$$Y = \beta_0 + \beta_1 X_1 + \alpha_r D + \epsilon.$$

In this case the parameter α_r will be:

$$\alpha_r = \{\alpha_0 P[D=1| X_1=0](1-P[D=1| X_1=0])P[X_1=0] \tag{8}$$

$$+ \alpha_1 P[D=1| X_1=1](1-P[D=1| X_1=1])P[X_1=1]\}/E\{P[D=1| X_1](1-P[D=1| X_1])\}.$$

Equation (8) means that, like $E[Y_1 - Y_0| D=1]$, the regression parameter α_r is a true weighted average of the underlying heterogeneous treatment effects, α_0 and α_1. Moreover, a feature common to both α_r and $E[Y_1 - Y_0| D=1]$ is that values of X where $P(D=1| X)$ is equal to either 1 or 0 are given zero weight. The difference between α_r and $E[Y_1 - Y_0| D=1]$ is solely in the nature of the weights at values of X where both veterans and nonveterans are observed. The weights underlying $E[Y_1 - Y_0| D=1]$ are proportional to the probability of veteran status at each value of the covariates. In contrast, the weights underlying α_r are proportional to the variance of veteran status at each value of the covariates. This means that $E[Y_1 - Y_0| D=1]$ weights each of the underlying treatment effects by $P[D=1| X]$ while α_r weights each of the underlying treatment effects by $P[D=1| X](1-P[D=1| X])$.

To see the importance of this difference, suppose, for example, that $P(X_1=1)=1/2$ and that $P[D=1| X_1=0]=.9$ while $P[D=1| X_1=1]=1/2$. Applying equations (7) and (8), it is possible to show that:

$$E[Y_1 - Y_0| D=1] = .64\alpha_0 + .36\alpha_1,$$

$$\alpha_r = .26\alpha_0 + .74\alpha_1.$$

Thus, while $E[Y_1 - Y_0| D=1]$ reflects the fact that veterans are much more likely to have $X_1=0$ in this numerical example, α_r puts more weight on the treatment effect for those with $X=1$ because the variance of D is larger for that group.

Note that the fact that the regression and matching estimates in Figure 9 diverge in later years could, in principle, be attributable either to changes in the weights underlying the two sets of estimates or to an increase in treatment effect heterogeneity over time. But since $P[D| X]$ does not vary from year to year in the data set used here, it must be the case that treatment effect heterogeneity is increasing over time. This heterogeneity may originate in differences in the long-term impact of military service on men with different levels of schooling and test scores. In particular, military service is probably most beneficial for men with the fewest civilian opportunities. Therefore, applicant screening in favor of those with higher scores and more schooling ends up admitting applicants least likely to benefit from military service. Unlike regression estimates, which weight heterogeneous effects by a conditional variance, $E[Y_1 - Y_0 | D=1]$ reflects the over-representation of low-benefit groups among veterans by producing a somewhat lower average treatment effect. Finally, note that while differences between $\hat{\alpha}_c$ and regression estimates are noticeable, the regression estimates generate similar conclusions regarding the overall effects of military service.

5. Estimation using the ASVAB misnorming

The estimates in Section 4 control for the major characteristics used by the military to screen applicants to the Armed Forces. But most of the applicants who do not enlist nevertheless qualify for enlistment (Berryman, Bell, and Lisowski 1983.) This raises the possibility that even among qualified applicants veterans could have earnings potential different from that of nonveterans. In this section estimates are presented from an IV strategy designed to control for unobserved differences between enlisting and non-enlisting applicants.

The approach taken here uses the fact that correction of errors in the AFQT score scale generated a sharp change in military admissions standards over the period 1979-81. A detailed history of the AFQT and other applicant screening tests since World War II is given in my earlier paper on military applicants (Angrist 1993a; see also Maier and Sims 1986 and Maier and Truss 1983). The importance of changes in the AFQT score scale for this project is that an applicant whose corrected AFQT score put him in category IV would have had a much higher chance of being accepted into the military if he had applied before 1980 than if he had applied after 1980. In contrast, applicants with AFQT scores in group III were virtually unaffected by correction of the AFQT score scale.

The impact of changes in the AFQT score scale on the probability of enlistment is illustrated in Table 6, which shows enlistment probabilities in the applicant population by AFQT group and year of application. As before, these probabilities were estimated from a sample limited to 1979-82 applicants with AFQT scores in groups III and IV who did not graduate college. Column 5 of the table shows that for applicants with AFQT scores in group IIIa (the highest-scoring group in the sample), the probability of enlistment was the same in 1979 and 1982. Similarly, the probability of enlistment for white men scoring in AFQT group IIIb fell by only about 5 percentage points from 1979-82. In contrast, the probability of enlistment for white men scoring in AFQT group IVc (the lowest scoring group in the sample) fell about 21 percentage points from 1979-82, with sharp drops in 1980 and 1981. The contrast for nonwhites is even larger: the probability of enlistment for nonwhite applicants scoring in AFQT group IIIa rose very slightly from 1979-82, while the probability of enlistment for nonwhite applicants scoring in AFQT group IVc fell by about 23 percentage points over the same period.

A dummy-variable regression is used to express the main features of Table 6 compactly while controlling for other covariates. Recall that for each race, the available covariates are year of birth, schooling, application year, and AFQT score. Let $m(X_j)$ be the jth row of a design matrix for a saturated model for schooling and year of birth. To simplify notation, the IV approach is initially developed using a model with a single effect for application year (early versus late) and a single effect for AFQT score group (high versus low). Let $a_j = 1$ denote applicants who applied in 1979-80 and let $s_j = 1$ to denote applicants with AFQT scores in group IV. Those with $a_j = 1$ are called "early" applicants and those with $s_j = 1$ are called "low scoring" applicants.

Consider the following saturated model for the probability of enlistment given X_j, a_j, and s_j:

$$E[v_j | X_j, a_j, s_j] = \delta_0 + m(X_j)\delta_1 + a_j\delta_2 + s_j\delta_3 + a_j m(X_j)\delta_{21} + s_j m(X_j)\delta_{31} + a_j s_j \delta_{23} + a_j s_j m(X_j)\delta_{231}.$$

In this expression, δ_1, δ_2, δ_3 are vectors of main effects for X_j, a_j, and s_j; δ_{21}, and δ_{31} are interaction terms for (X_j, a_j) and (X_j, s_j); δ_{23} is an early/low-scoring interaction term, and δ_{231} is a vector of third-order effects that allows the early/low-scoring interaction term to differ at different values of X_j. This equation captures all the variation in the conditional probability of enlistment given X_j, a_j, and s_j.

To simplify the exposition, third-order terms are initially ignored. The conditional expectation of veteran status is therefore assumed to be:

$$(9) \qquad E[v_j | X_j, a_j, s_j] = \delta_0 + m(X_j)\delta_1 + a_j\delta_2 + s_j\delta_3 + a_j m(X_j)\delta_{21} + s_j m(X_j)\delta_{31} + a_j s_j \delta_{23}.$$

Differences-in-differences contrasts based on the rows and columns of Table 6 can be interpreted as estimates of interaction terms like δ_{23} in equation (9) when the additional covariates (X_j) are ignored. The coefficient δ_{23} describes how much more (on average and conditional on X_j) the probability of enlistment declined for low-scoring applicants between 1979-80 and 1981-82 than for high-scoring applicants. The value of δ_{23} for whites with non-missing earnings in 1991 is .142 and the value of δ_{23} for nonwhites with non-missing earnings in 1991 is .173. These estimates were computed from a weighted least-squares fit of the population $E[v_j | X_j, a_j, s_j]$ to equation (9).

5.1 Difference-in-differences estimates

A differences-in-differences estimator of the effects of military service on earnings matches early/low-scoring interaction terms in earnings functions to estimates of δ_{23} in equation (9). This approach is developed using the following model:

(10) $\bar{y}_{jt} = \beta_{0t} + m(X_j)\beta_{1t} + a_j\beta_{2t} + s_j\beta_{3t} + a_jm(X_j)\beta_{21t} + s_jm(X_j)\beta_{31t} + \alpha_{jt}v_j + \bar{\epsilon}_{jt},$

where the error term is assumed to satisfy $E[\bar{\epsilon}_{jt} | X_j, a_j, s_j]=0$. The link between equation (10) and the notation of Section 4.1 is as follows. Let Y_{0t} denote the potential earnings of nonveterans in year t and let Y_{1t} denote the potential earnings of veterans in year t. Then, we have:

$$E[Y_{0t} | X_j, a_j, s_j] = \beta_{0t} + m(X_j)\beta_{1t} + a_j\beta_{2t} + s_j\beta_{3t} + a_jm(X_j)\beta_{21t} + s_jm(X_j)\beta_{31t},$$

and

$$E[Y_{1t} - Y_{0t} | X_j, a_j, s_j] = \alpha_{jt}.$$

Thus, the regression function in (10) describes what veterans would have earned had they not served in the military, while α_{jt} is a heterogeneous treatment effect that varies across cells and years.

A key feature of equation (10) is that it includes no interaction terms in application year and AFQT score group. In contrast, the reduced form for veteran status, (9), includes a_js_j with coefficient δ_{23}. Estimation of equation (9) shows that such interaction terms are substantial, shifting the probability of enlistment by 14-18 percentage points. Thus, a_js_j is a potential instrumental variable. The strategy for IV estimation using a_js_j as an instrument takes the average effect of veteran status to be mean-independent of covariates, so that

(11) $E[\alpha_{jt} | X_j, A_j, S_j; v_j=1] = \alpha_t.$

The reduced form earnings equation in this constant-treatment-effect case is therefore:

(12) $E[\bar{y}_{jt} | X_j, a_j, s_j] = \pi_{0t} + m(X_j)\pi_{1t} + a_j\pi_{2t} + s_j\pi_{3t} + a_jm(X_j)\pi_{21t} + s_jm(X_j)\pi_{31t} + a_js_j\pi_{23t},$

where π_{23t} is equal to $\alpha_t\delta_{23}$. The ratio of π_{23t} to an estimate of δ_{23} from equation (9) is a covariates-adjusted differences-in-differences estimate of α_t that controls for secular application-year and AFQT score-group effects, as well as effects for X_j and interactions of X_j with a_j and s_j. Note, however, that omission of the third order term, $a_js_jm(X_j)$, from equation (9) means that equation (12) will be mis-specified if (9) is mis-specified. Because equations (9) and (12) are used primarily as an illustration, this problem is ignored here but corrected in the construction of IV estimates, below.

Weighted least squares estimates of π_{23t} in equation (12) are plotted in Figures 10a and 10b, along with 95 percent confidence bands. The estimates and standard errors were computed using the reciprocal of the variance of mean earnings conditional on X_j, A_j, and S_j as weights. Further details of the estimation procedure are given in the statistical appendix. Recall that δ_{23} is .142 for whites and .173 for nonwhites. Estimates of α_t for whites can therefore be obtained by multiplying estimates of π_{23t} by 7; Estimates of α_t for nonwhites can be obtained by multiplying estimates of π_{23t} by about 5.8.

For white applicants, the pattern of estimates of π_{23t} has a number of important similarities to the pattern of OLS and matching estimates found in previous figures. First, the estimates of π_{23t} are insignificantly different from zero through 1978. Second, the estimates of π_{23t} increase in 1979 and are significant and positive from 1981-83 when most of the veterans in the sample were in the military. From 1984 on, the estimates for whites are insignificantly different from zero. Estimates for nonwhites are also zero through 1978, with a sharp dip in 1980 and 1981. Like the estimates for whites, the estimates for nonwhites are significant and positive in 1982 and 1983. After 1983, however, the estimates of π_{23t} for nonwhites are negative and significant. For both groups, the positive estimates of π_{23t} are consistent with a treatment effect of between 1000 and 2500 dollars.

The interpretation of π_{23t} as the product of α_t and δ_{23} depends partly on the exclusion of an early/low-scoring interaction term from the underlying earnings function. This is a functional form restriction, the validity of which may turn on the details of model specification. For example, the absence of interaction terms in a model for levels does not imply the absence of interactions terms in a model for logs. Figures 11a and 11b therefore present estimates of π_{23t} for log earnings comparable to those for the level of earnings in Figures 10a and 10b. The log estimates of π_{23t} show a pattern similar to the pattern of estimates for levels. The log estimates for whites are positive and significant from 1980-84 while the log estimates for nonwhites are positive and significant from 1981-85. Log estimates for both groups in later years are either negative or insignificant.

5.2 IV estimates

IV estimates were calculated using a saturated first-stage equation given all covariates in the data set other than veteran status. Let $m(A_j)$, and $m(S_j)$ denote the jth rows of design matrices for application year and AFQT score group. The first-stage equation is:

$$\begin{aligned}(13)\quad E[v_j \mid X_j, A_j, S_j] = \;& \delta_0 + m(X_j)\delta_1 + m(A_j)\delta_2 + m(S_j)\delta_3 + \\ & [m(A_j)\otimes m(X_j)]\delta_{21} + [m(S_j)\otimes m(X_j)]\delta_{31} + [m(A_j)\otimes m(S_j)]\delta_{23} + \\ & [m(A_j)\otimes m(S_j)\otimes m(X_j)]\delta_{231},\end{aligned}$$

where the coefficients (δs) now denote conformable vectors of main effects and second and third-order interaction terms. Equation (13) has two sources of variation not found in equation (9). First, $a_j s_j$ in equation (9) is replaced with a full set of application-year and AFQT score-group interaction terms. Second, a full set of third-order terms, $[m(A_j)\otimes m(S_j)\otimes m(X_j)]$, allow the application-year and AFQT score-group interaction terms to differ at each value of X_j.

The interpretation of third-order terms is that for some values of X_j, the application-year and AFQT score-group interaction terms are likely to be more important than for others. For example, the early/low-scoring interaction term is about 5 percentage points higher for whites without a high school diploma than for white high school graduates. This is because AFQT scores are relatively unimportant in admissions decisions for applicants with a high school diploma. Note that because equation (13) is saturated, any least squares fit will simply recover the empirical probability of enlistment given A_j, S_j, and X_j.

Figures 12a and 12b plot IV estimates of α_t in the following equation:

$$\text{(14)}\quad E[\bar{y}_{jt} \mid X_j, A_j, S_j] = \beta_{0t} + m(X_j)\beta_{1t} + m(A_j)\beta_{2t} + m(S_j)\beta_{3t} + [m(A_j)\otimes m(X_j)]\beta_{21t} + [m(S_j)\otimes m(X_j)]\beta_{31t} + \alpha_t E[v_j \mid A_j, S_j, X_j],$$

using (13) as the model for $E[v_j \mid A_j, S_j, X_j]$. The estimates were computed by replacing $E[v_j \mid A_j, S_j, X_j]$ with the corresponding empirical population probability and replacing $E[\bar{y}_{jt} \mid X_j, A_j, S_j]$ with the corresponding estimate of the population mean. This equation is identified by exclusion of $[m(A_j)\otimes m(S_j)]$ and $[m(A_j)\otimes m(S_j)\otimes m(X_j)]$. Provided these exclusion restrictions are valid, equation (14) accurately captures the relationship between $E[\bar{y}_{jt} \mid X_j, A_j, S_j]$ and $E[v_j \mid A_j, S_j, X_j]$. Additional details on the estimation procedure are given in the statistical appendix.

To facilitate comparisons with the earlier results, Table 7 reports the matching estimates from Figures 4a and 4b in columns 1-2 and the IV estimates from Figures 12a and 12b in columns 3-4. The IV estimates for whites have a number of important similarities to the matching estimates. Both approaches show significant positive effects in the early 1980s, although the IV estimates for 1980-83 are much larger than the matching estimates. The matching estimates for the 1984-90 earnings of whites are negative and statistically significant. The IV estimates are also mostly negative after 1984, although they are not significantly different from zero. But confidence bands for the IV estimates in later years bracket the OLS and matching estimates.

The IV estimates for nonwhites also show substantial positive treatment effects in the early 1980s. Moreover, IV estimates for the 1979-84 earnings of nonwhites are remarkably similar to the matching estimates. But the IV estimates for earnings after 1985 differ from the matching estimates in that the IV estimates are negative while the matching estimates are positive and significant. It should be noted, however, that none of the negative IV estimates for earnings after 1984 are significantly different from zero. Finally, note that unlike the reduced form estimates of equation (12) for nonwhites, the results in Figure 12b show no evidence of a sharp dip in treatment effects in 1980-81.

5.3 IV estimates in models with treatment effect heterogeneity

The behavioral interpretation of the IV estimates in Figures 12a and 12b turns not only on functional form considerations but also on the assumption that the treatment effect,

$E[\alpha_{jt} | X_j, A_j, S_j; v_j=1]$, does not actually vary with A_j and S_j.[16] The assumption that the treatment effect is fixed is almost certainly violated in transition years when some applicants are entering or leaving the military. For example, the treatment effect on applicants who have not yet entered the military is necessarily zero, while the effect on applicants already in the military need not be zero. Similar differences in treatment effects for earnings in later years are generated by the fact that applicants who entered the military at different times leave the military at different times. Recall that the matching estimates in Figures 5a and 5b show clear differences in treatment effects by application year corresponding to a staggered pattern of discharge from the military.

It is easy to see the implications of treatment-effect heterogeneity in an example with one excluded instrument as in equations (9-12). Suppose that for α_{jt} in equation (10), we have,

(15) $E[\alpha_{jt} | X_j, A_j, S_j; v_j=1] = \alpha_{0t}(1-a_j) + \alpha_{1t}a_j.$

In equations for 1980 earnings, for example, α_{0t} (the effect for late applicants) should be zero but α_{1t} (the effect for early applicants) is probably not zero. The knowledge that one treatment effect is zero provides potentially useful identifying information. In general, however, under specification (15) the parameter π_{23t} in equation (12) is equal to

$$\tilde{\pi}_t \equiv \alpha_{0t}(\delta_{23}-\delta_3) + \alpha_{1t}\delta_3.$$

[16]Variation in $E[\alpha_{jt} | X_j, A_j, S_j; v_j=1]$ with X_j is less problematic than variation with A_j and S_j. In an IV context, any treatment effect heterogeneity from exogenous covariates that are not used as instruments is reflected in a variance-weighted average of the type discussed in Section 4.3, above. See Angrist and Imbens (1995) for details.

Without additional restrictions, neither α_{0t} or α_{1t} can be recovered from estimates of $\tilde{\pi}_t$.[17] Moreover, $\tilde{\pi}_t$ is not a convex combination of α_{0t} and α_{1t}.

The main implication of this discussion is that estimates of π_{23t} for the 1979-81 earnings of applicants should not be interpreted as being proportional to an underlying average treatment effect. Estimates for later years when large numbers of veterans are leaving the military could also be problematic. Recall that the matching estimates in Figures 5a and 5b show a distinct pattern of effects by application year. These differences correspond to staggered discharges for men who enlisted at different times. The resulting variation in treatment effects will tend to confound IV estimates based on the exclusion of early/low-ability interaction terms.

The problem of treatment effect heterogeneity is solved as follows. For purposes of estimation, suppose that the model for $E[\alpha_{jt} | X_j, A_j, S_j]$ allows for a heterogeneous treatment effect involving a_j and s_j:

(16) $E[\alpha_{jt} | A_j, S_j, X_j; v_j=1] = \alpha_{0t} + \alpha_{2t}a_j + \alpha_{3t}s_j$.

Combining (16) and (10), we have

(17) $$E[\bar{y}_{jt} | X_j, A_j, S_j] = \beta_{0t} + m(X_j)\beta_{1t} + m(A_j)\beta_{2t} + m(S_j)\beta_{3t} + [m(A_j)\otimes m(X_j)]\beta_{21t} + [m(S_j)\otimes m(X_j)]\beta_{31t} + [\alpha_{0t}p_j + \alpha_{2t}a_jp_j + \alpha_{3t}s_jp_j],$$

where $p_j \equiv E[v_j | A_j, S_j, X_j]$. The three-parameter treatment effect in equation (17) is identified because the terms excluded from equation (10) include not only early/low-scoring interactions but also the full set of (A_j, S_j) interaction terms and third-order terms for $(A_j, S_j,$

[17]The expression for $\tilde{\pi}_t$ is obtained by substituting (15) into (10), taking expectations conditional on X, A, and S, and simplifying.

X_j). The exclusion of these terms introduces enough variability in p_j to identify coefficients on interaction terms in the treatment effect, a_jp_j and s_jp_j.

Figures 13a and 13b plot the results of weighted least squares estimation of equation (17) with a first-stage for veteran status specified as in equation (13). To make the results as comparable as possible to those in Section 4, the figures show an "effect of treatment on the treated" computed by weighting estimates of α_{0t}, α_{2t}, and α_{3t} by the proportion of early and low-scoring applicants among veterans. For example, Figure 13a for whites plots estimates of $[\alpha_{0t} + .50\alpha_{2t} + .27\alpha_{3t}]$ because 50 percent of white veterans applied early and 27 percent were low-scoring. The estimates and underlying standard errors are also plotted in columns 6-7 of Table 7. Note that results from a more general model specifying the heterogeneous treatment effect as

$$E[\alpha_{jt} |\ A_j, S_j, X_j; v_j=1] = \alpha_{0t} + \alpha_{2t}a_j + \alpha_{3t}s_j + \alpha_{23t}a_js_j$$

are virtually identical to those reported in the figures.

For whites, the main difference between Figures 13a and 12a are the larger estimates for 1981-83 earnings in Figure 13a. Figure 13a shows an especially large outlying estimate of over 5000 dollars generated by the heterogeneous treatment effects model for 1982 earnings. Figure 13a also shows more positive estimates than Figure 12a for earnings after 1983. As in Figure 12a, however, these estimates are not significant. Figure 13b, which plots the average heterogeneous treatment effect for nonwhites, also differs from Figure 12b in that the estimates for 1980-84 are larger. As with the matching results for nonwhites, estimates from the heterogeneous treatment effects model after 1984 suggest a positive long-run impact of military service on the earnings of nonwhites. Although these effects decline from 1983-1989, they remain statistically significant through 1987. The 1991 estimate of nearly $1400 is significant as well.

With the exception of results for 1981 and 1982, IV results from the heterogeneous treatment effects model for nonwhites are remarkably similar to the matching estimates. But an important ambiguity is the conflict between the IV results from the constant treatment effects model and the IV results from the heterogeneous treatment effects model. To provide some guidance in choosing between these models, columns 5 and 7 of Table 7 report chi-square goodness-of-fit statistics that measure the fit of each model to mean earnings. This test statistic is the minimized objective function used to produce the IV estimates. Under the null hypothesis that the models fit the means, the chi-square statistics have degrees of freedom equal to the number of cell averages ($E[\bar{y}_{jt} | X_j, A_j, S_j]$) minus the number of free parameters in the model.

The estimates in column 3 were produced by a model that imposes two additional restrictions beyond those imposed to produce the estimates in column 6. The difference in the chi-square statistics in columns 5 and 8 should therefore have two degrees of freedom when these two additional restrictions are satisfied. The .95 critical value of a $\chi^2(2)$ is about 6. A comparison of columns 5 and 8 in Table 7 suggests that, with the exception of the results for 1980-82 earnings, there is little difference in fit between the two underlying models for the earnings of whites. For nonwhites, however, the contrast in goodness-of-fit statistics favors the unrestricted model allowing for treatment effect heterogeneity in 1979-82 and for every year after 1983. It is the results from this model that are positive and closest to the matching estimates for nonwhites.

6. Conclusions

Even in the population of applicants to the military, simple comparisons of earnings by veteran status exaggerate the positive effects of military service and mask possible negative effects. Matching and OLS estimates that control for age, schooling, year of application, and test scores show that in the early 1980s white applicants who entered the military did in fact earn more than comparable applicants who did not serve. But estimates of the effect of military service on white applicants after 1984 are negative or close to zero. Matching and OLS results for nonwhites show large positive veteran effects through 1984, with smaller positive effects through the end of the sample period in 1991.

Results of IV estimation are generally supportive of the matching and OLS results, although the IV estimates suggest that the matching estimates underestimate the gains to veterans while they were in the military. As with the matching estimates, IV estimates of longer-term effects on whites are negative or statistically indistinguishable from zero. For nonwhites, one set of IV estimates suggests that there are no long-run earnings benefits from military service. But results from a better-fitting model that allows for treatment effect heterogeneity echo the matching and OLS findings, showing modest positive effects through the end of the sample period. The finding of long-run civilian earnings benefits for nonwhites contrasts with earlier negative or ambiguous results for veterans of the AVF (Lawrence, Ramsberger, and Gribben 1989). On the other hand, the results reported here for positive earnings contradict findings from the NLSY showing benefits of military service for working veterans (Bryant, Samaranayake, and Wilhite 1993, Magnum and Ball 1989.)

Although the effects on military service on civilian earnings appear to be negative for whites and modest for nonwhites, military service clearly raised the earnings of both groups of veterans while they were in the military. It should be noted, however, that part of this beneficial effect may be attributable to the fact that the period of service for most of the veterans studied here overlapped with a cyclical downturn in the US economy. The results presented here should therefore be used cautiously in attempts to predict the earnings impact of military service on future cohorts of veterans.

Finally, note that for both racial groups, any positive veteran effects on civilian earnings appears to be largely attributable to improved employment prospects for veterans. Conditional on working, matching estimates for the 1986-90 earnings of nonwhites are negative or insignificant while estimates for the 1984-91 earnings of whites show a large penalty. These findings suggest that efforts to moderate the impact of military downsizing should be focused on raising the probability of employment for minority youth who would otherwise have found positions in the military.

References

Angrist, J.D., "Using the Draft Lottery to Measure the Effect of Military Service on Civilian Labor Market Outcomes," *Research in Labor Economics*, Volume 10, Ronald Ehrenberg, ed., Greenwich: JAI Press, Inc., 1989.

__________, "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence From Social Security Administrative Records," *American Economic Review*, 80 (June 1990): 313-335.

__________, "The Draft Lottery and Voluntary Enlistment in the Vietnam Era," *Journal of the American Statistical Association* 86(415) (September 1991): 584-595.

__________, "The Misnorming of the U.S. Military's Entrance Examination and its Effect on Minority Enlistments," University of Wisconsin, Institute for Research on Poverty, Discussion paper No. 1017-93, September 1993a.

__________, "The Effect of Veterans Benefits on Education and Earnings," *Industrial and Labor Relations Review*, July 1993b.

__________, "Conditioning on the Probability of Selection to Control Selection Bias," NBER Technical Working Paper, June 1995.

__________, and G. Imbens, "Sources of Identifying Information in Evaluation Models", NBER Technical Working Paper No. 117, December 1991.

__________, and G. Imbens, "Two-Stage Least Squares Estimates of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association* 89, June 1995.

__________, and Alan B. Krueger, "Why Do World War Two Veterans Earn More Than Nonveterans?," *Journal of Labor Economics*, January 1994: 74-97.

Barnow, B., G. Cain, and A. Goldberger, "Selection on Observables," *Evaluation Studies Review Annual*, 5, 43-59, Beverly Hills: Sage, 1981.

Berger, M.C., and B.T. Hirsch, "The Civilian Earnings Experience of Vietnam-Era Veterans," *Journal of Human Resources* 18 (Fall 1983): 455-79.

Berryman, S.E, R.M. Bell, and W. Lisowski, *The Military Enlistment Process: What Happens and Can it be Improved?*, Rand Report R-2986-MRAL, Santa Monica, May 1983.

Beusse, W.E., "The Impact of Military service on Low Aptitude Men," Alexandria, VA: Air Force Human Resources Laboratory, February 1974.

Bryant, R.B., V.A. Samaranayake, and A. Wilhite, "The Effect of Military Service on the Subsequent Civilian Wage of the Post-Vietnam Veteran," *Quarterly Review of Economics and Finance*, January 1993: 15-30.

Card, D.E., and Daniel Sullivan, "Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment," *Econometrica* 56 (January 1988): 497-530.

Cooper, R.V.L., *Military Manpower and the All-Volunteer Force*, Report R-1450-ARPA, Santa Monica: RAND Corporation, September 1977.

Cutright, Phillips, "The Civilian Earnings of White and Black Draftees and Nonveterans," *American Sociological Review*, 39 (June 1974): 317-27.

Defense Department, *Population Representation in the Military Services*, Fiscal year 1987, Office of the Assistant Secretary of Defense (Force Management and Personnel), Washington, D.C., August 1988.

Defense Department, *Population Representation in the Military Services*, Fiscal year 1991, Office of the Assistant Secretary of Defense (Force Management and Personnel), Washington, D.C., October 1992.

Dehejia, R.H., and S. Wahba, "Causal Effects in Nonexperimental Studies: Re-evaluating the Evaluation of Training Programs," Harvard Economics Department, May 1995.

Eitelberg, Mark J., J. Lawrence, B. Waters, and L. Perelman, *Screening For Service: Aptitude and Education Criteria for Military Entry*, Alexandria, Human Resources Research Organization, September 1984.

Heckman, J.J. and R. 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.

Holland, P., and D. Rubin, *Test Equating*, New York: Academic Press, 1982.

Imbens, G.W., and J. D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62 (March 1994): 467-476.

Laurence, J.H., "Crew Cuts: Effects of the Defense Drawdown on Minorities," Chapter 2 in *Military Cutbacks and the Expanding Role of Education*, Office of Educational research and Improvement, US department of Education, October 1992.

__________, P.F. Ramsberger, and M.A. Gribben, "Effects of Military Experience on the Post-Service Lives of Low-Aptitude Recruits: Project 100,00 and the ASVAB Misnorming," FR-PRD-89-29, Alexandria, Va: Human Resources Research Organization, December 1989.

Magnum, S.L., and D.E. Ball, "The Transferability of Military-Provided Occupational Training in the Post-Draft Era," *Industrial and Labor Relations Review* 42(2), (January 1989): 230-245.

Maier, M.H., and W.H. Sims, *The ASVAB Score Scales: 1980 and World War II*, Alexandria, Va.: Center for Naval Analyses, Report No. 116, 1986.

__________, and A.R. Truss, *Original Scaling of ASVAB Forms 5/6/7: What Went Wrong?*, Alexandria, Va.: Center for Naval Analyses, Research Contribution No. 457, 1983.

Newey, W.K., "Efficient Instrumental Variables Estimation of Nonlinear Models," *Econometrica* 58, 1990, 809-837.

Orvis, B.R, and M.T. Gahart, *Enlistment Among Applicants for Military Service*, Rand Report R-3359-FMP, Santa Monica, January 1990.

Rosenbaum, P.R., and D.B. Rubin (1983), "Assessing Sensitivity to and Unobserved Binary Covariate in an Observational Study with Binary Outcomes," *Journal of the Royal Statistical Society Series B*, 45(2), 212-218.

Rubin, D.B. (1974), "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies," *Journal of Educational Psychology*, 66, 688-701.

__________, (1977), "Assignment to a Treatment Group on the Basis of a Covariate", *Journal of Educational Statistics*, 2, 1-26.

U.S. Congress, "Social Representation in the U.S. Military," Congressional Budget Office, Washington DC, USGPO, October 1989.

U.S. Department of Health and Human Services, Social Security Administration, Office of Research and Statistics, *Annual Statistical Supplement to the Social Security Bulletin*, Washington DC: January 1993.

Table 1: Male Applicant Population (thousands)

Race	Year Applied 1976	1977	1978	1979	1980	1981	1982
White	339.5	286.9	235.9	253.1	348.6	387.3	309.8
Percent veteran	53	52	54	55	53	49	52
Nonwhite	128.6	114.8	103.6	119.5	134.3	149.3	112.5
Percent veteran	44	46	50	46	41	36	43

Notes: The population is the same as that described in Angrist (1993a, Table 4) excluding those with less than a 9th grade education at the time of application. Veterans are applicants identified as entrants to the military in the year of application or in one of the two years following application.

Table 2: Applicant Sample (thousands)

Race	Year Applied 1976	1977	1978	1979	1980	1981	1982
White	49.2	46.5	40.0	39.4	52.9	57.9	47.3
Percent veteran	56	53	55	57	54	50	53
Nonwhite	50.9	48.1	44.6	51.9	57.0	63.7	48.7
Percent veteran	49	49	52	49	44	38	45

Notes: Approximately 90 percent of this sample is a self-weighting random sample from the population in Table 1, stratified by race. Observations in small cells are over-sampled. Details of the sampling scheme are provided in the data appendix.

Table 3: AFQT Score Distribution

Race	Application year	AFQT Score Group							
		I	II	IIIa	IIIb	IVa	IVb	IVc	V
White									
	1976	5.4	27.2	15.0	19.6	11.5	7.9	7.5	5.8
	1977	4.4	25.1	16.9	18.8	12.7	8.5	8.4	5.2
	1978	4.4	25.1	17.3	19.2	12.7	8.6	8.5	4.3
	1979	3.8	24.4	16.9	20.0	12.9	8.9	8.1	5.0
	1980	3.9	25.7	17.7	20.7	12.6	7.9	7.2	4.3
	1981	2.9	27.7	17.1	21.4	12.8	7.3	6.7	4.2
	1982	3.4	30.9	18.1	22.3	12.0	6.1	5.0	2.2
Nonwhite									
	1976	0.3	5.5	7.2	16.9	15.4	14.0	19.4	21.4
	1977	0.0	5.0	7.1	15.4	16.0	15.6	22.1	18.9
	1978	0.0	5.3	7.8	16.1	16.2	15.0	22.2	17.5
	1979	0.1	4.9	7.1	15.3	16.9	16.7	22.1	17.1
	1980	0.1	5.5	7.3	15.6	17.0	16.3	21.3	16.9
	1981	0.2	6.6	8.1	18.7	18.5	14.7	18.5	14.8
	1982	0.1	8.8	10.1	23.0	21.6	14.6	14.3	7.6

Notes: The table shows unweighted frequency distributions for applicants successfully matched to Social Security earnings data. AFQT group I includes percentile scores 93-100, AFQT group II includes percentile scores 65-92, AFQT group IIIa includes percentile scores 50-64, AFQT group IIIb includes percentile scores 31-49, AFQT group IVa includes percentile scores 21-30, AFQT group IVb includes percentile scores 16-20, AFQT group IVc includes percentile scores 10-15, and AFQT group V denotes the lower decile in the reference population.

Table 4: Veteran and Non-veteran Earnings

	Whites			Nonwhites		
Year	Mean (1)	Difference (2)	s.e. (3)	Mean (4)	Difference (5)	s.e. (6)
A. All earnings						
74	182.7	-26.1	7.0	157.2	-4.9	4.4
75	237.9	-41.4	6.3	216.9	-0.6	4.5
76	473.4	-47.9	8.1	413.6	-14.5	6.4
77	1012.9	-7.1	11.3	820.9	-13.0	9.1
78	2147.1	40.3	16.7	1677.9	58.1	13.4
79	3560.7	188.0	21.0	2797.0	340.3	16.2
80	4709.0	572.9	23.4	3932.2	1154.3	18.0
81	6226.0	855.5	27.2	5218.8	1920.0	20.7
82	7200.6	1508.5	30.3	6150.2	2917.1	23.4
83	8398.1	1390.5	34.4	7221.1	2889.9	27.0
84	9874.2	652.8	39.5	8377.2	2202.9	30.5
85	10972.7	469.8	44.6	9306.8	1955.5	34.4
86	12004.5	543.7	50.4	10106.5	1881.3	38.7
87	13045.7	663.9	54.6	10833.0	2050.1	41.8
88	14136.1	904.3	58.3	11480.1	2175.0	44.9
89	14716.1	1169.1	61.0	11751.4	2379.1	47.6
90	14886.1	1300.8	63.0	11904.3	2483.6	49.4
91	14407.9	1559.6	64.6	11518.7	2758.8	50.8
B. Positive earnings						
74	1546.6	-24.8	48.1	1457.3	32.2	33.4
75	1755.1	0.1	38.1	1596.9	-20.4	27.1
76	2275.1	-13.7	32.7	2003.4	-20.2	26.2
77	2915.6	-24.3	28.1	2498.5	6.6	24.3
78	3949.6	20.3	27.2	3336.6	62.3	23.7
79	4988.4	91.5	26.5	4197.4	307.5	22.1
80	5799.2	358.7	26.8	5186.2	981.1	21.8
81	7034.5	408.8	29.0	6350.0	1446.6	23.2
82	8004.8	722.0	31.7	7358.8	1913.0	25.6
83	9227.6	561.4	35.2	8453.7	1807.1	28.7
84	10708.6	-59.4	39.8	9530.7	1301.0	31.6
85	11938.5	-178.5	44.7	10533.4	1099.2	35.3
86	13177.6	-106.2	50.6	11537.4	1014.0	39.7
87	14364.4	21.6	54.8	12409.6	1180.1	43.0
88	15580.1	326.1	58.3	13213.1	1316.0	46.3
89	16360.9	572.8	60.9	13735.0	1518.4	49.4
90	16761.9	678.7	63.1	14199.3	1608.4	51.6
91	16677.7	891.0	65.5	14302.4	1844.6	54.3

Notes: The table shows average earnings and the difference in earnings by veteran status for white and nonwhite 1979-82 applicants with AFQT scores in groups III and IV. This sample includes 128,968 white applicants and 175,262 nonwhite applicants, about half of whom served in the military. Panel A of the table shows average earnings including zeros and Panel B shows average positive earnings. The fraction with zero earnings ranges from 86 percent for whites and 89 percent for nonwhites in 1974 to 7.9 percent for whites and 11.5 percent for nonwhites in 1985. Standard errors for the veteran/nonveteran contrast are given in columns 3 and 6. Earnings are in 1991 dollars, deflated by the CPI.

Table 5: Veteran/nonveteran Contrast Adjusted for all Covariates ($\hat{\alpha}_c$)

Year	Contrast for whites	s.e.	Contrast for nonwhites	s.e.
A. All earnings				
74	-10.3	7.2	-25.7	5.0
75	-27.4	6.4	-16.4	5.4
76	-36.2	8.3	-47.9	7.6
77	-15.6	12.0	-72.2	10.9
78	-34.4	18.0	-51.5	16.0
79	-10.5	23.2	121.0	20.1
80	180.1	27.1	741.6	23.4
81	249.5	32.4	1299.9	28.2
82	783.3	36.4	2186.0	32.0
83	588.8	41.1	2103.8	36.7
84	-235.7	46.9	1333.0	41.4
85	-521.3	52.6	932.3	46.2
86	-557.3	59.0	720.9	51.2
87	-548.0	63.9	751.0	55.2
88	-415.5	68.2	708.2	59.5
89	-248.6	71.2	799.7	62.9
90	-154.5	73.6	824.9	65.4
91	29.8	75.6	1026.1	67.2
B. Positive earnings				
74	-58.2	54.8	55.0	35.9
75	-27.8	42.1	-51.6	31.8
76	-15.7	34.8	-106.6	30.9
77	-6.7	29.6	-107.2	28.9
78	-26.5	29.2	-84.3	28.0
79	-94.2	29.5	36.3	27.5
80	-60.2	31.1	473.5	28.7
81	-212.8	34.6	739.8	32.0
82	102.1	37.6	1216.3	34.8
83	-136.6	41.6	1077.8	38.8
84	-878.8	46.9	451.5	42.8
85	-1117.3	52.5	93.9	47.3
86	-1166.1	59.0	-123.8	52.5
87	-1151.4	63.8	-118.0	56.6
88	-975.2	67.9	-175.3	61.1
89	-835.4	70.7	-104.5	65.0
90	-769.3	73.3	-108.0	67.9
91	-581.9	76.1	69.5	71.3

Notes: The table reports the differences in earnings by veteran status plotted in Figures 4 and 6. The differences are averaged within-cell contrasts weighted by population veteran counts to produce an estimate of the effect of veteran status on veterans. The sample is the same as that underlying Table 4.

Table 6: Misnorming and the probability of enlistment

Race	Year of Application	AFQT Group IVc	IVb	IVa	IIIb	IIIa
		(1)	(2)	(3)	(4)	(5)
whites	1979	.264	.517	.591	.623	.631
	1980	.145	.428	.515	.594	.623
	1981	.033	.226	.359	.590	.625
	1982	.054	.232	.360	.571	.631
nonwhites	1979	.296	.580	.659	.681	.685
	1980	.198	.506	.602	.665	.677
	1981	.044	.301	.470	.668	.684
	1982	.063	.287	.480	.665	.686

Notes: The table shows the probability of enlistment by AFQT score group and year of application. Estimates are for the population of 1979-82 non-college graduate applicants with AFQT scores in groups III and IV. This population includes 849,983 white applicants and 403,686 nonwhite applicants.

Table 7: Comparison of Estimates

Year	Matching estimates		IV estimates					
	$E[Y_1 - Y_0 \mid D=1]$		Constant t.e.			Heterogeneous t.e.		
	Estimates (1)	s.e. (2)	Estimate (3)	s.e. (4)	χ^2(dof) (5)	Estimate (6)	s.e. (7)	χ^2(dof) (8)
A. Whites								
74	-10.3	7.2	2.5	10.3	93(93)	11.6	15.9	92(91)
75	-27.4	6.4	-1.2	6.0	137(156)	-5.0	6.8	132(154)
76	-36.2	8.3	-5.8	8.3	189(211)	-36.6	17.1	183(209)
77	-15.6	12.0	16.2	28.3	243(252)	-18.4	35.5	239(250)
78	-34.4	18.0	204.9	80.7	289(266)	164.0	126.5	289(264)
79	-10.5	23.2	186.1	183.3	330(268)	-190.8	276.6	326(266)
80	180.1	27.1	560.9	250.5	409(268)	681.4	412.2	395(266)
81	249.5	32.4	1727.8	305.7	370(268)	2417.2	550.3	351(266)
82	783.3	36.4	2893.3	350.3	422(268)	5094.0	611.4	402(266)
83	588.8	41.1	1652.1	408.3	314(268)	1864.4	753.2	308(266)
84	-235.7	46.9	153.9	467.0	257(268)	-184.0	885.1	256(266)
85	-521.3	52.6	-788.7	517.1	279(268)	-1137.2	994.2	279(266)
86	-557.3	59.0	-595.7	575.3	238(268)	1310.3	1114.5	234(266)
87	-548.0	63.9	12.4	613.5	266(268)	1826.1	1198.7	262(266)
88	-415.5	68.2	-4.6	650.2	275(268)	372.9	1279.8	274(266)
89	-248.6	71.2	-368.2	676.5	248(268)	-37.5	1330.7	248(266)
90	-154.5	73.6	-995.6	694.6	252(268)	-160.4	1364.3	250(266)
91	29.8	75.6	-776.6	700.9	268(268)	549.6	1370.3	264(266)
B. Nonwhites								
74	-25.7	5.0	1.1	6.5	102(113)	-8.3	11.6	101(111)
75	-16.4	5.4	0.6	5.8	187(177)	-4.4	6.9	180(175)
76	-47.9	7.6	0.9	9.0	204(222)	-19.0	16.9	199(220)
77	-72.2	10.9	-22.0	30.5	259(243)	-31.1	33.6	257(241)
78	-51.5	16.0	42.2	62.2	273(256)	39.1	70.3	271(254)
79	121.0	20.1	150.3	126.7	377(257)	484.3	173.8	364(255)
80	741.6	23.4	637.1	168.0	725(257)	1899.5	238.1	511(255)
81	1299.9	28.2	1689.2	202.1	725(257)	4036.7	294.5	410(255)
82	2186.0	32.0	3144.8	226.1	715(257)	4816.3	327.3	645(255)
83	2103.8	36.7	2217.3	262.8	239(257)	2492.7	383.3	236(255)
84	1333.0	41.4	957.3	298.3	267(257)	1927.4	429.9	257(255)
85	932.3	46.2	2.0	335.6	286(257)	1154.2	488.2	270(255)
86	720.9	51.2	-685.6	376.3	311(257)	1008.4	548.6	283(255)
87	751.0	55.2	-454.7	403.7	297(257)	1448.0	601.3	270(255)
88	708.2	59.5	-381.3	432.0	292(257)	773.0	638.2	282(255)
89	799.7	62.9	-456.7	450.0	312(257)	642.5	661.9	303(255)
90	824.9	65.4	-578.5	465.7	315(257)	1049.3	686.1	291(255)
91	1026.1	67.2	-612.3	473.7	308(257)	1362.0	692.3	280(255)

Notes: Matching estimates are from Panel A in Table 5 and the controlled contrast in Figures 4a and 4b. The first set of IV estimates are the same as those reported in Figures 12a and 12b. The second set of IV estimates are the same as those reported in Figures 13a and 13b.

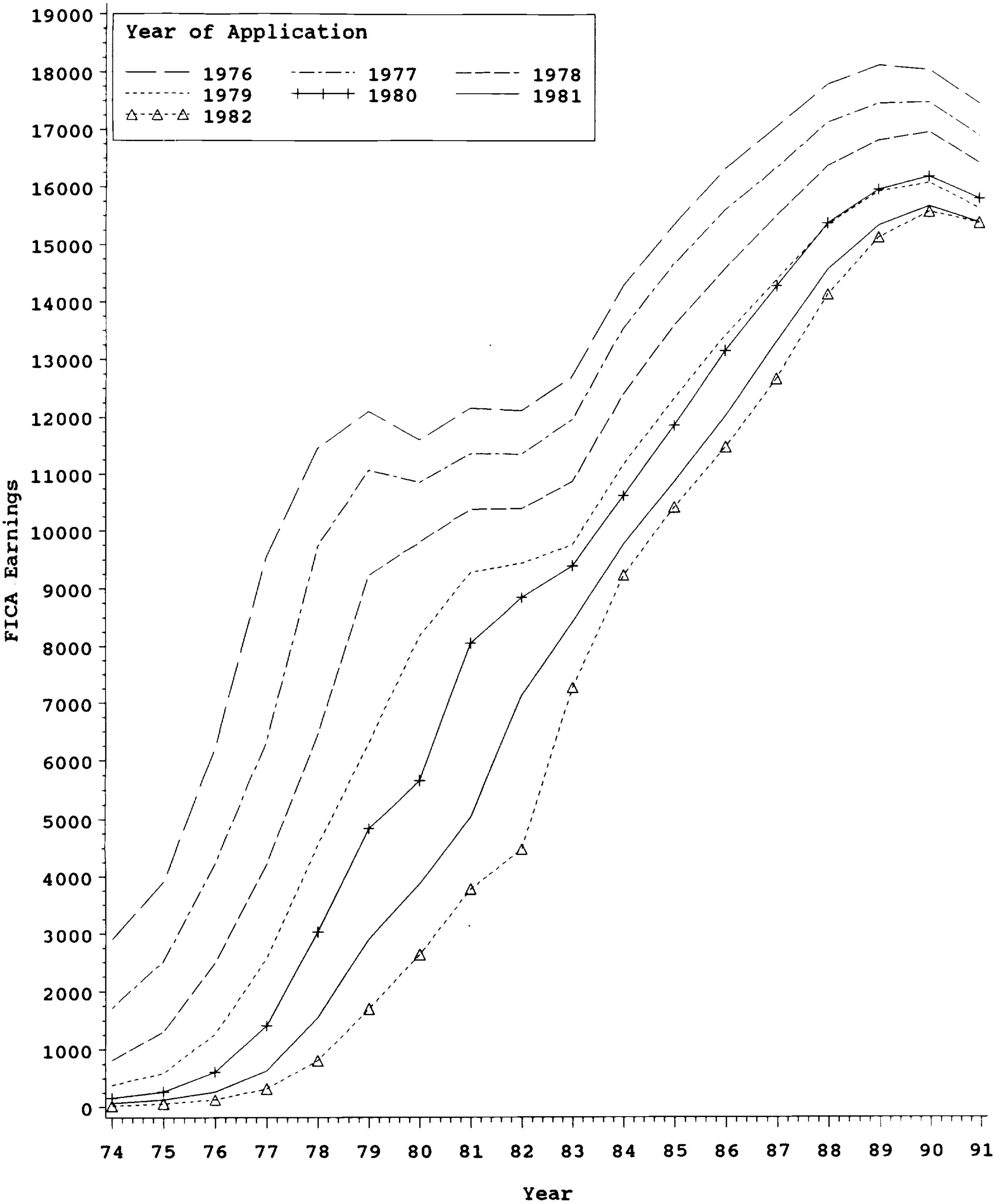

Figure 1a. Real Social Security earnings.
White applicants to the military in 1976-82.

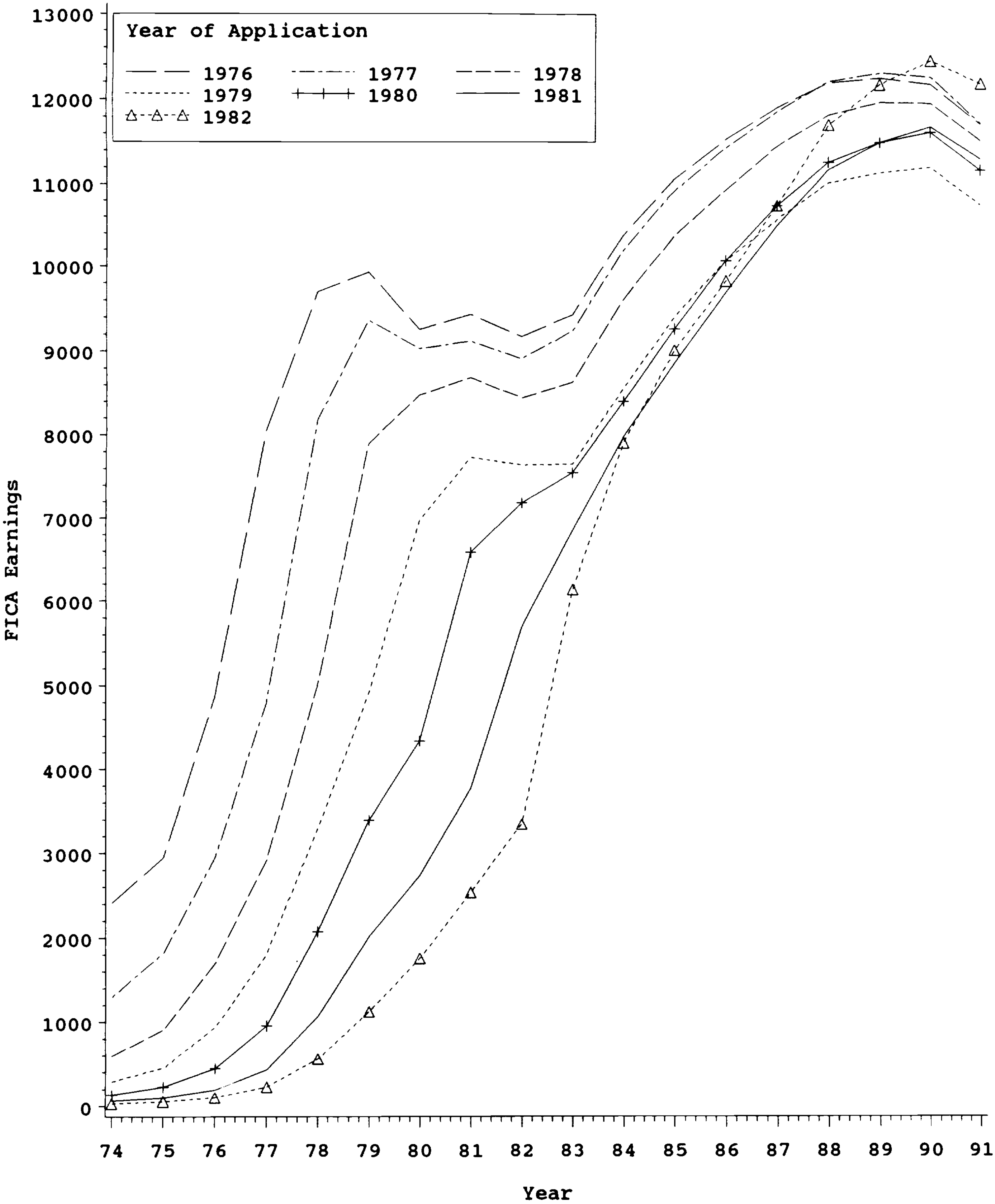

Figure 1b. Real Social Security earnings.
Nonwhite applicants to the military in 1976-82.

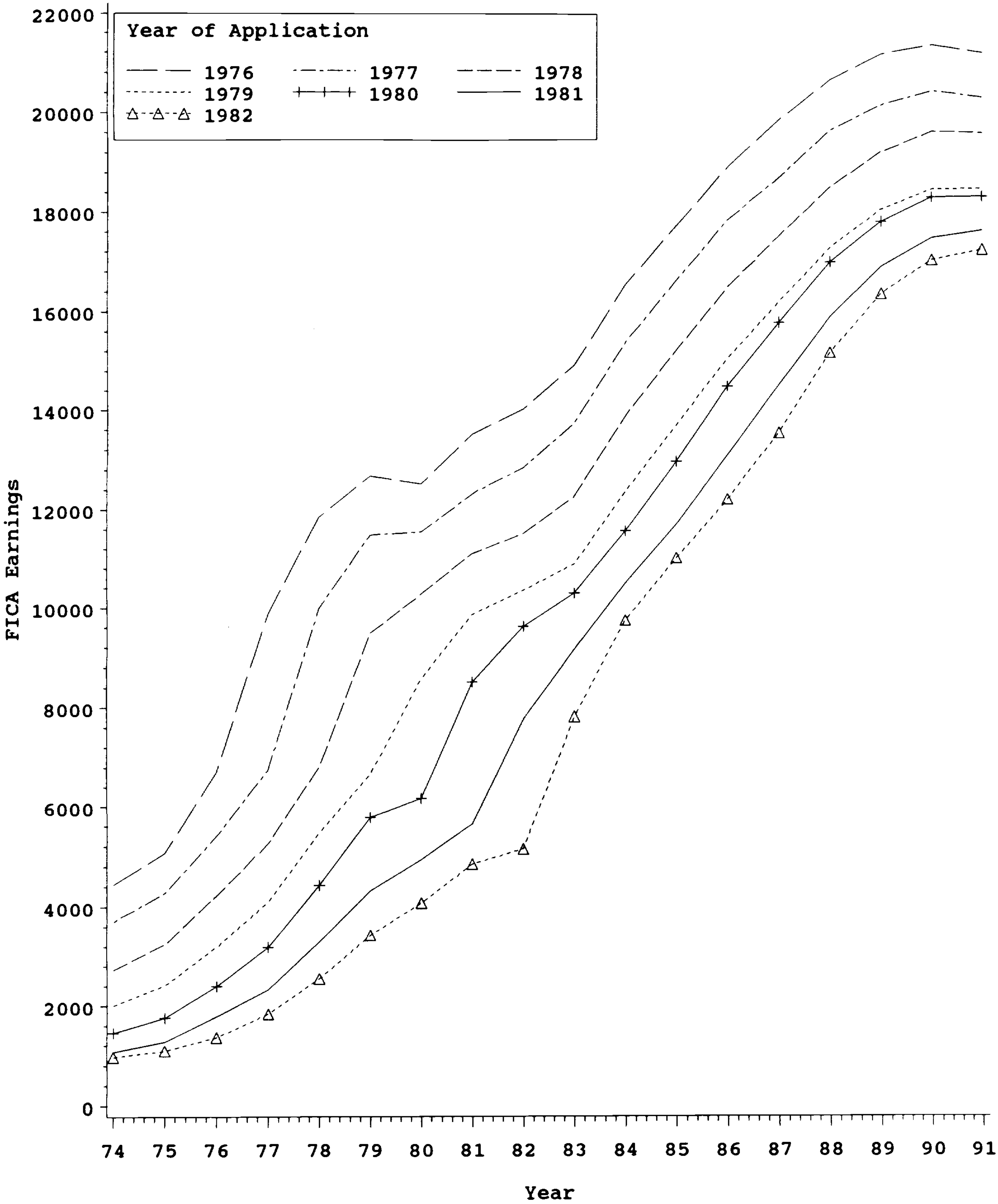

Figure 2a. Real positive Social Security earnings.
White applicants to the military in 1976-82.

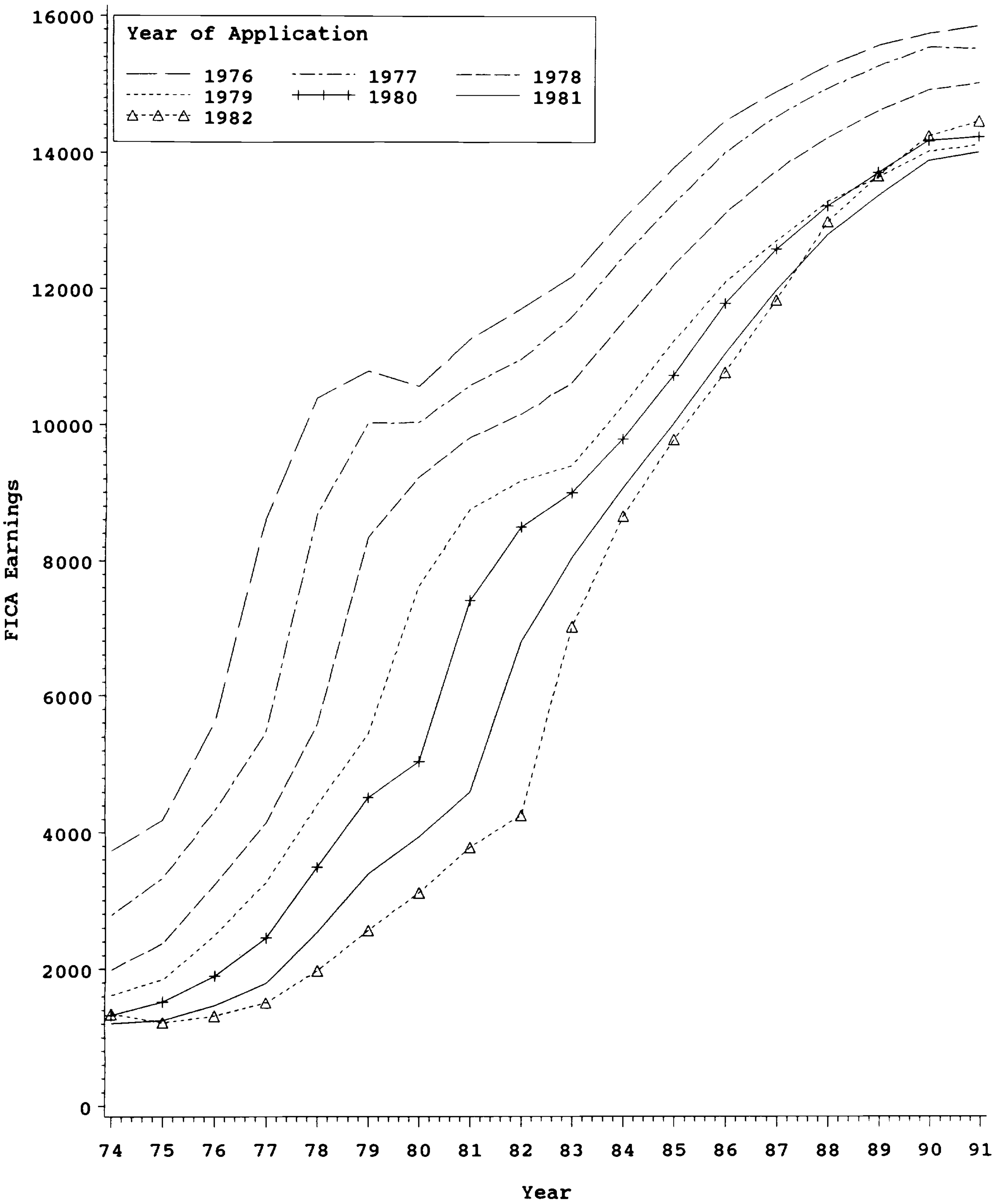

Figure 2b. Real positive Social Security earnings.
Nonwhite applicants to the military in 1976-82.

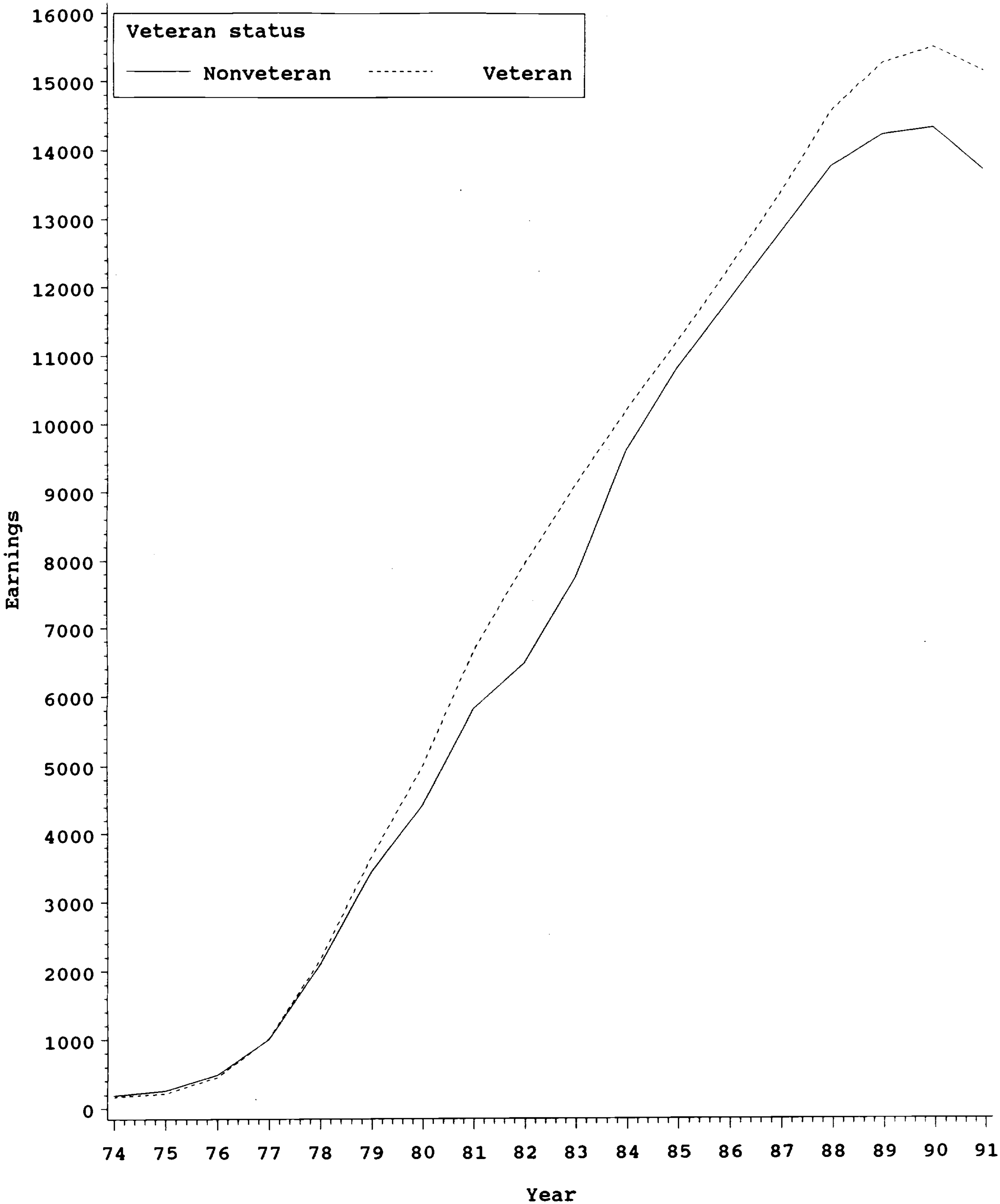

Figure 3a. Real Social Security earnings.
White applicants in AFQT groups III and IV, who applied in 1979-82.

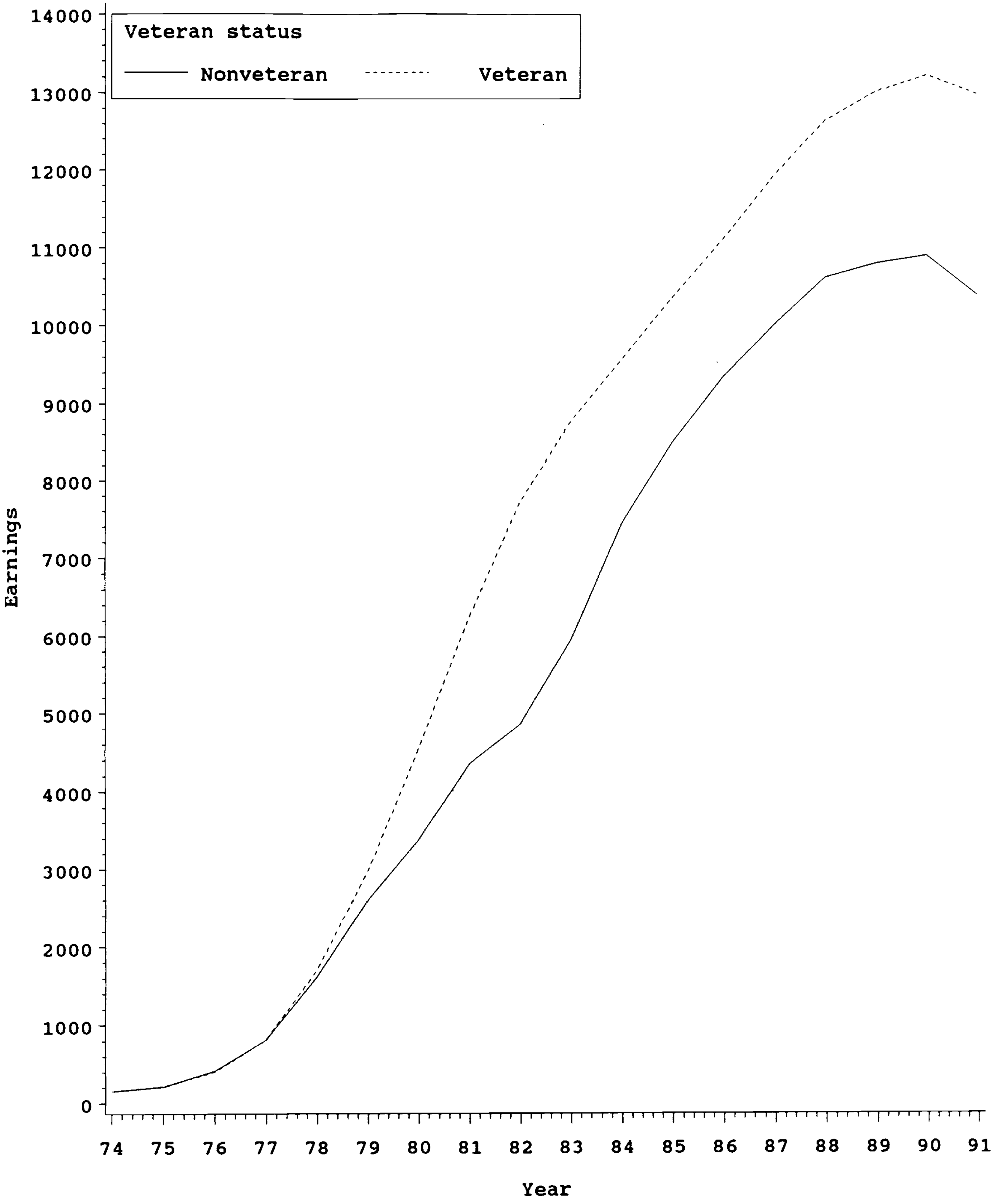

Figure 3b. Real Social Security earnings.
Nonwhite applicants in AFQT groups III and IV, who applied in 1979-82.

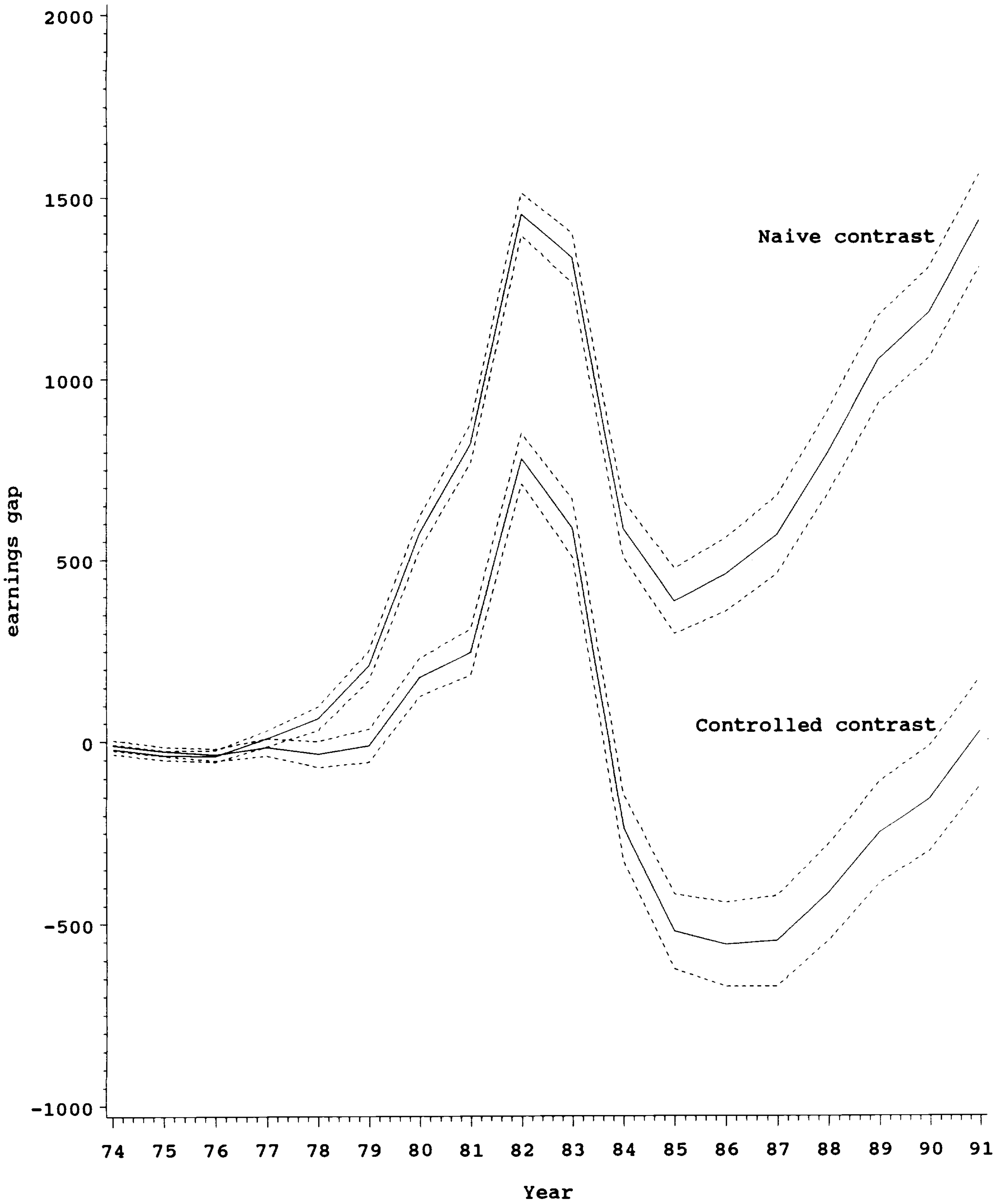

Figure 4a. Veteran/non-veteran earnings gap and confidence bands for Whites. Applicants in AFQT groups III and IV, who applied in 1979-82.

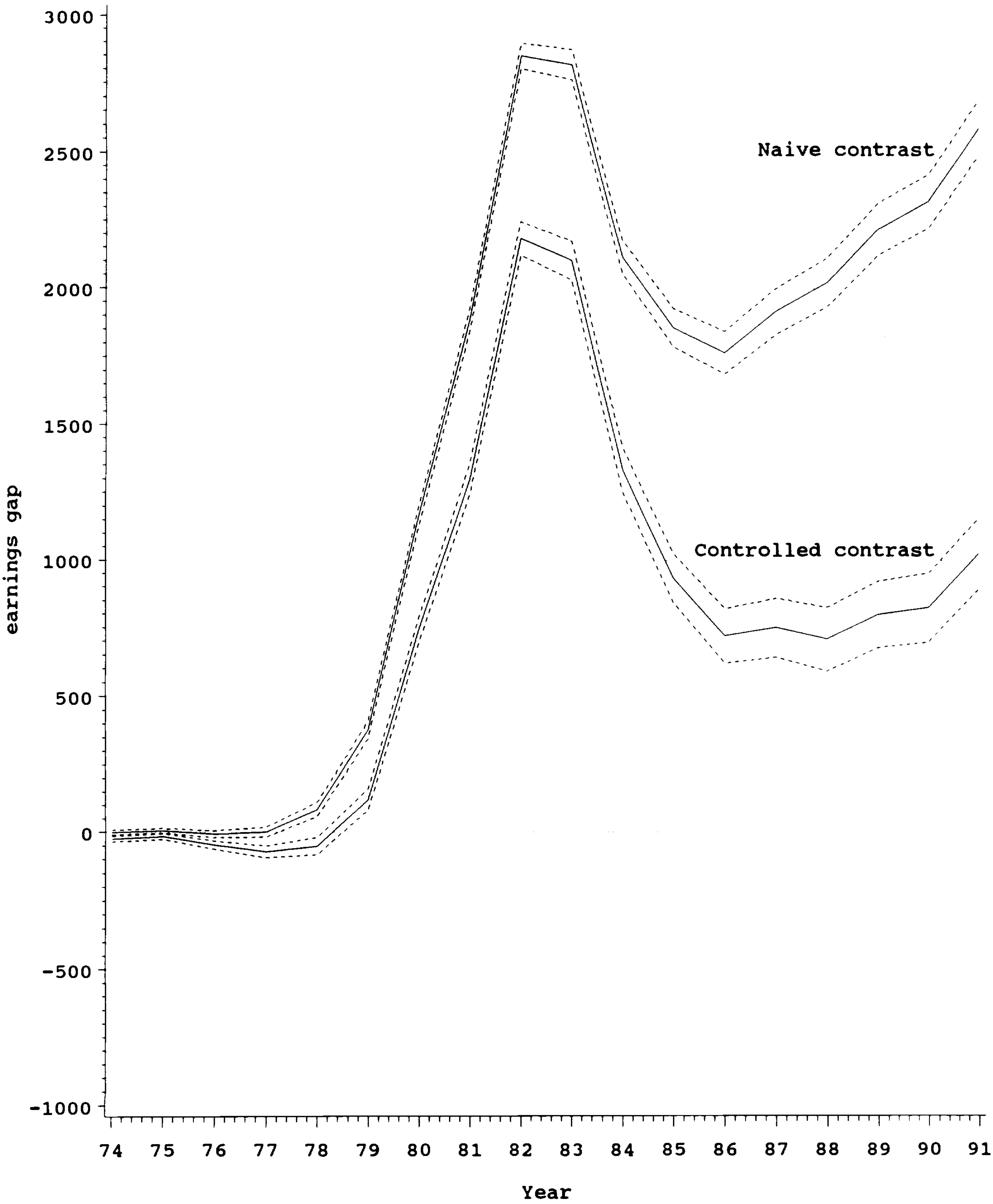

Figure 4b. Veteran/non-veteran earnings gap and confidence bands for Nonwhites. Applicants in AFQT groups III and IV, who applied in 1979-82.

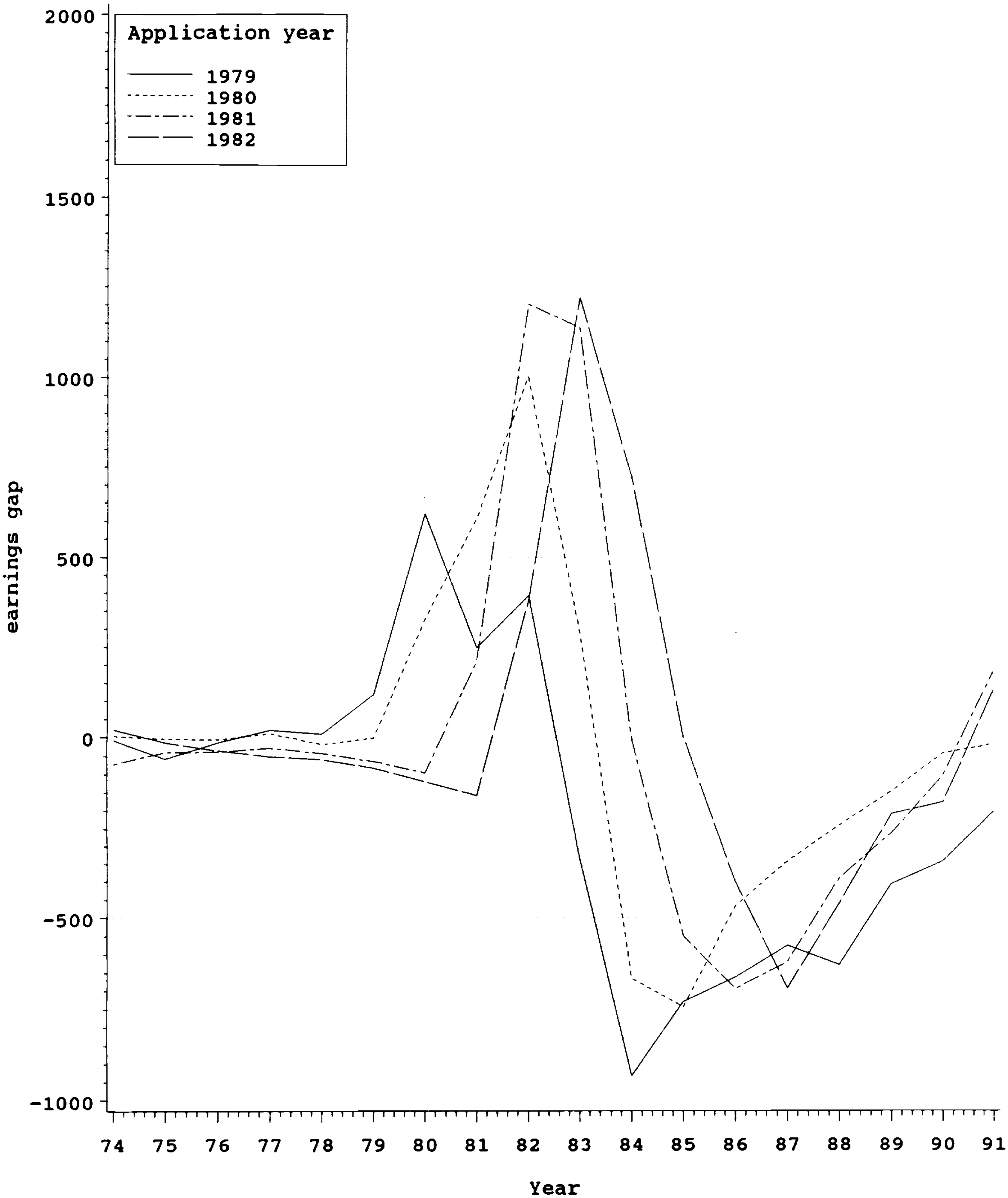

Figure 5a. Veteran/non-veteran earnings gap for Whites by application year. Applicants in AFQT groups III and IV, who applied in 1979-82. Contrasts in the figure control for all observable veteran/nonveteran differences.

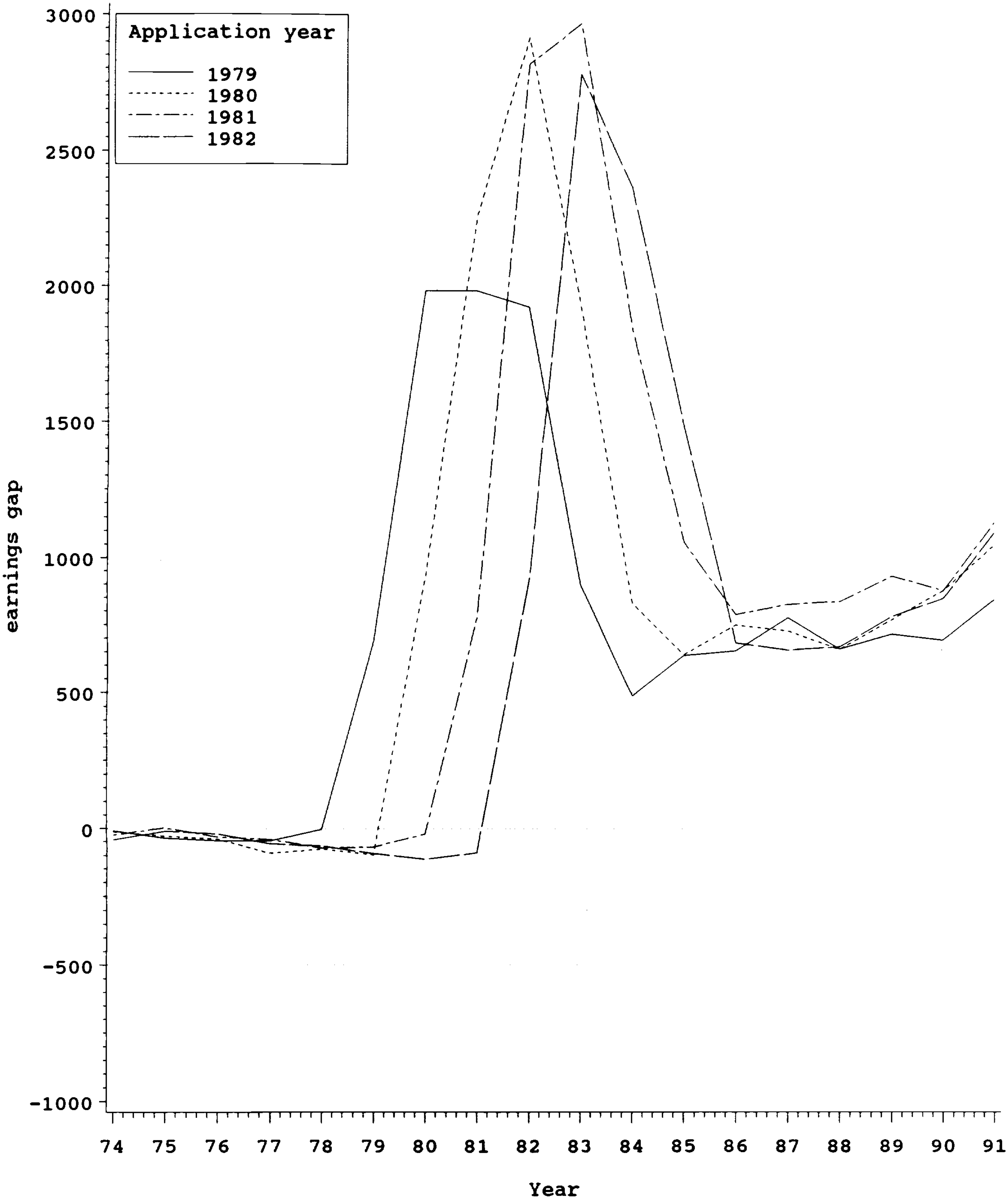

Figure 5b. Veteran/non-veteran earnings gap for Nonwhites by application year. Applicants in AFQT groups III and IV, who applied in 1979-82. Contrasts in the figure control for all observable veteran/nonveteran differences.

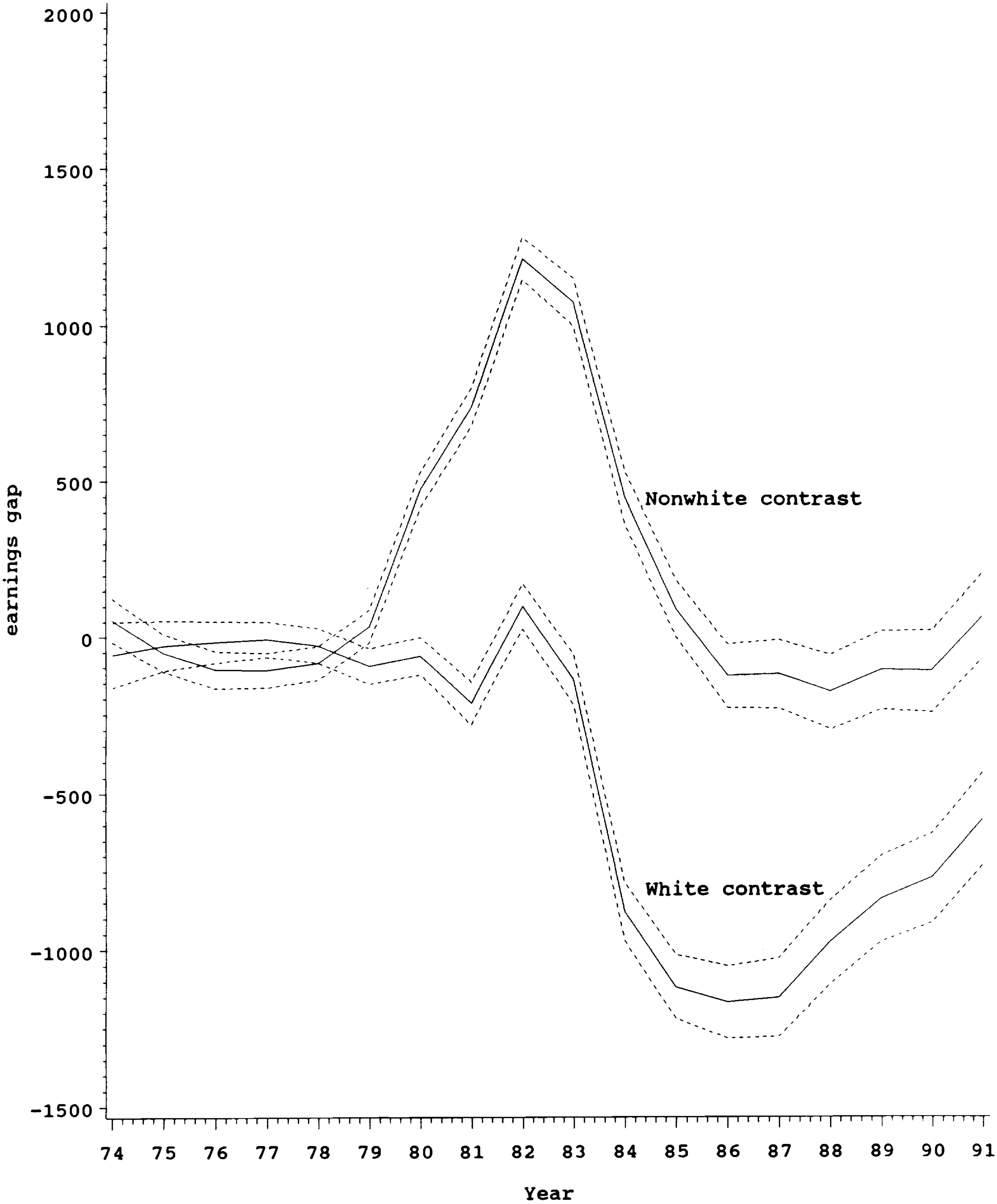

Figure 6. Veteran/non-veteran POSITIVE earnings gap and confidence bands for controlled contrast. 1979-82 applicants in AFQT groups III and IV.

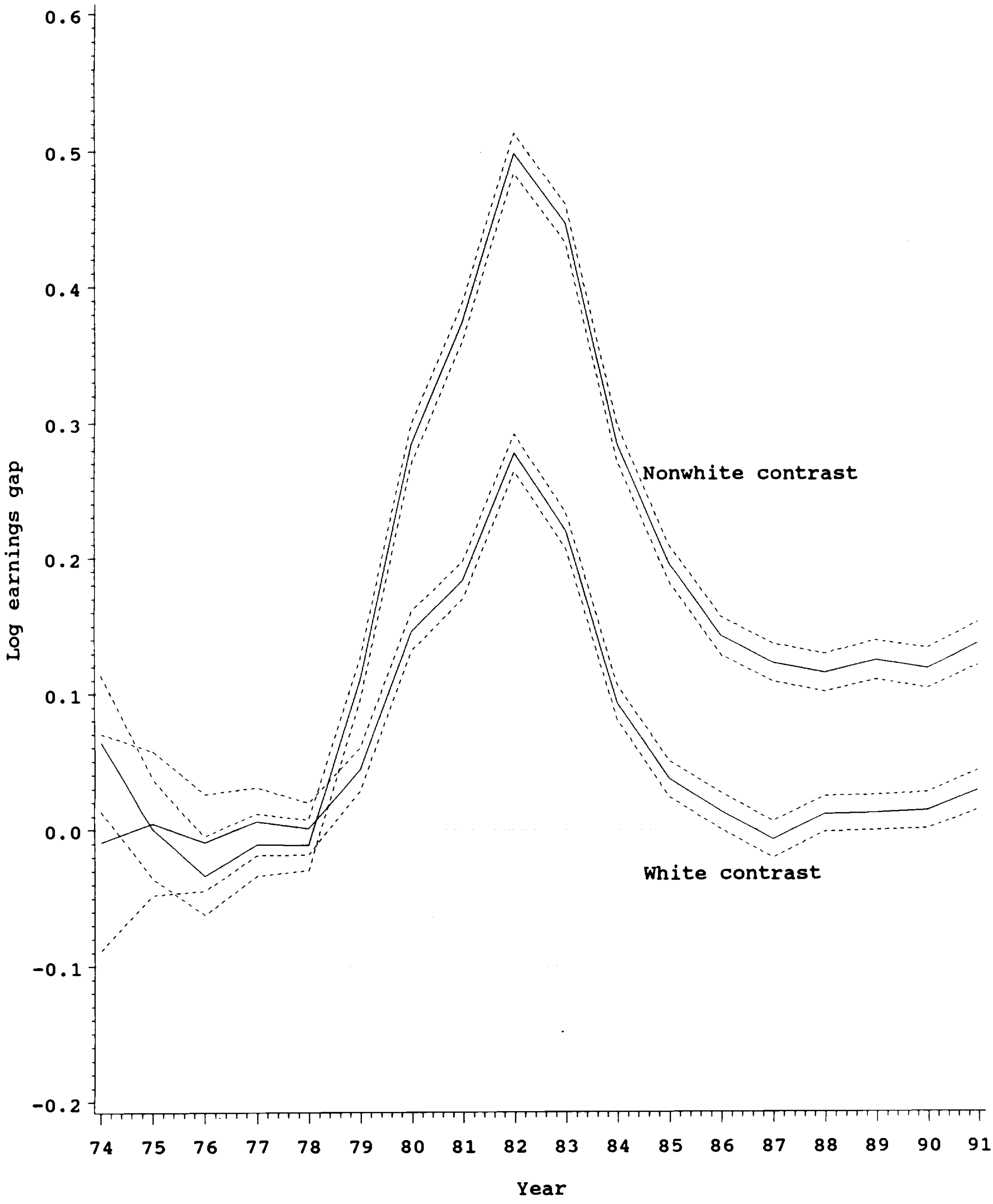

Figure 7. Veteran/non-veteran LOG earnings gap and confidence bands for controlled contrast. 1979-82 applicants in AFQT groups III and IV.

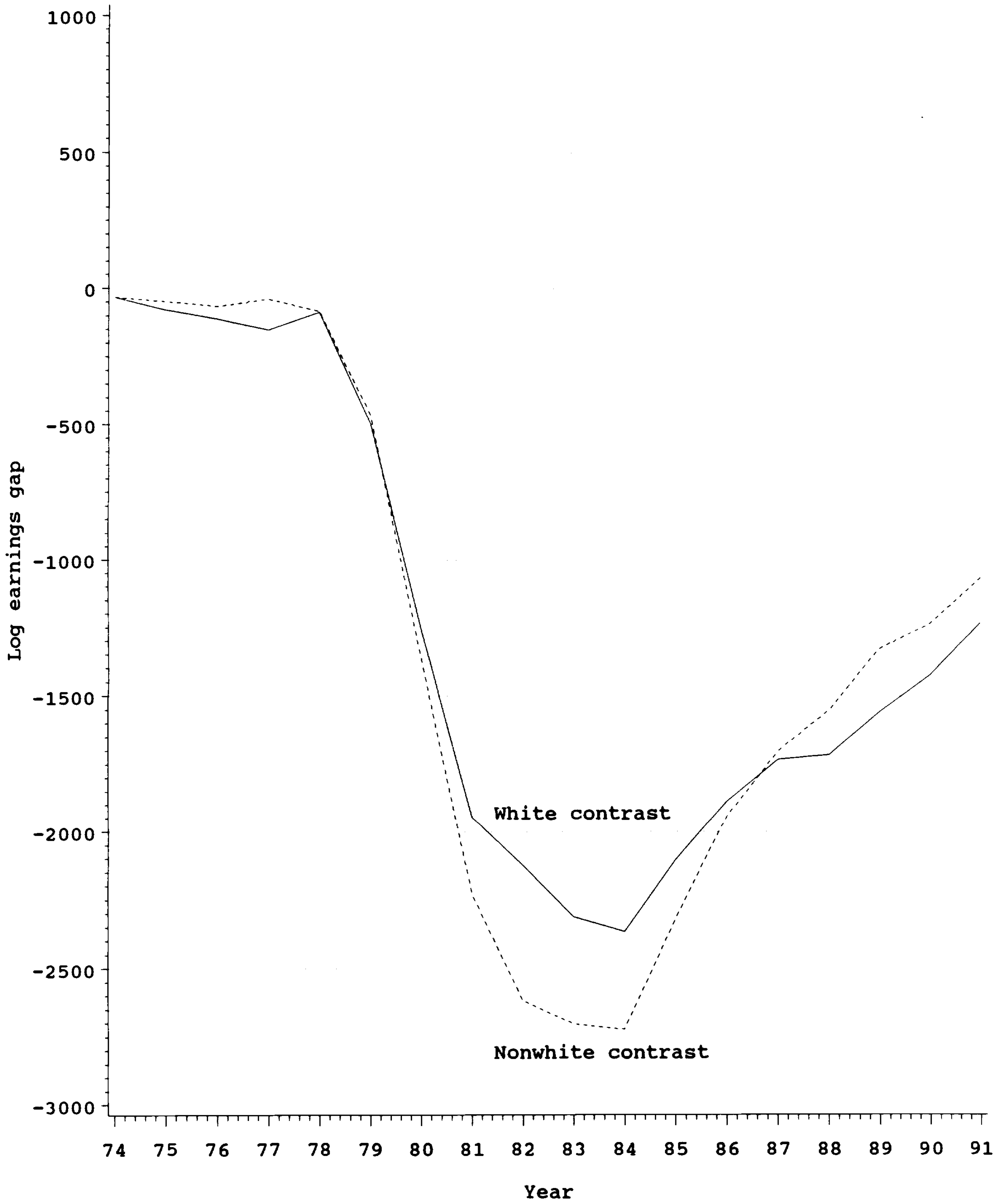

Figure 8. Veteran/non-veteran standard deviation gap for controlled contrast. 1979-82 applicants in AFQT groups III and IV.

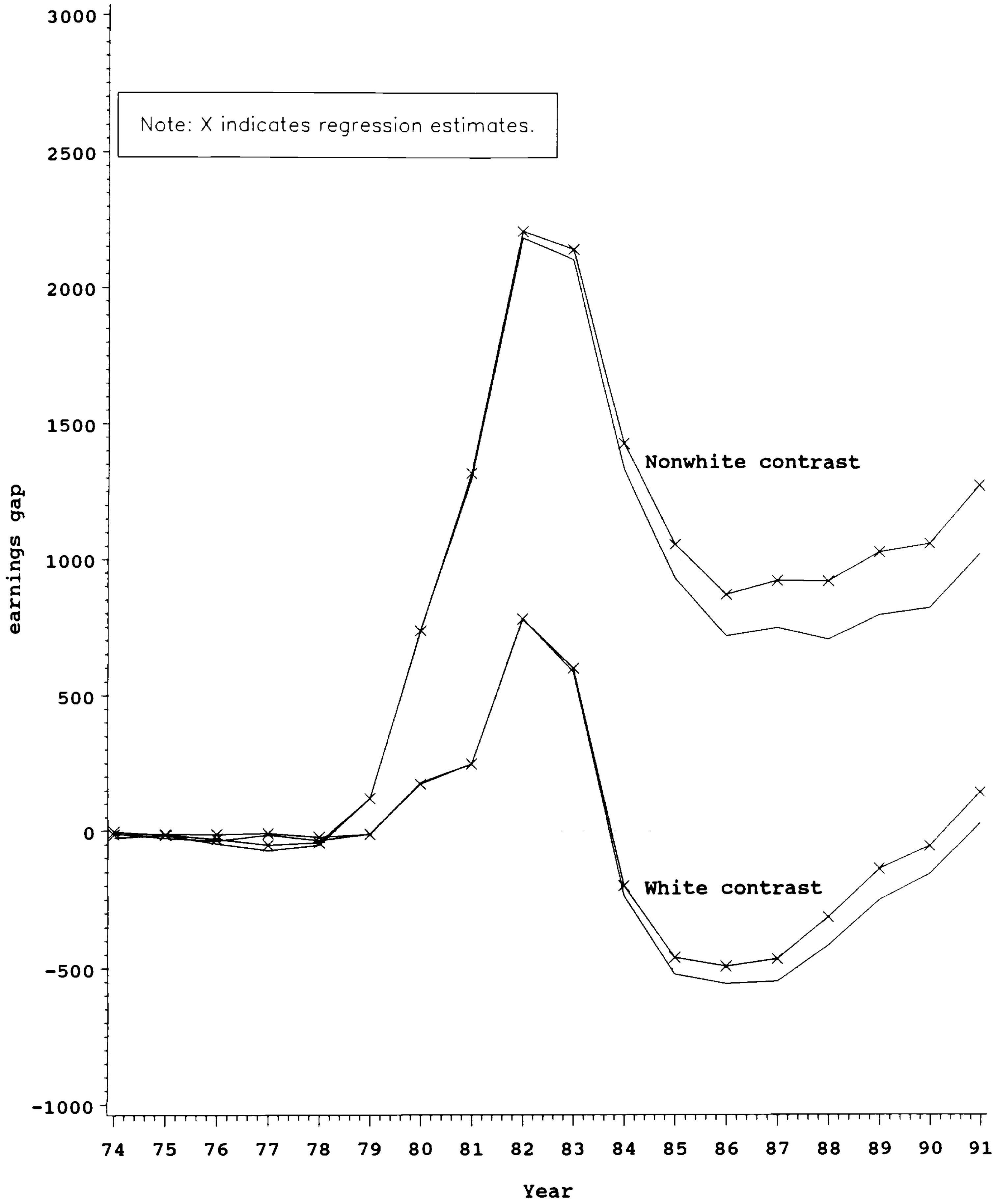

Figure 9. Veteran/non-veteran earnings gap: weighted contrasts and regression estimates. 1979-82 applicants in AFQT groups III and IV.

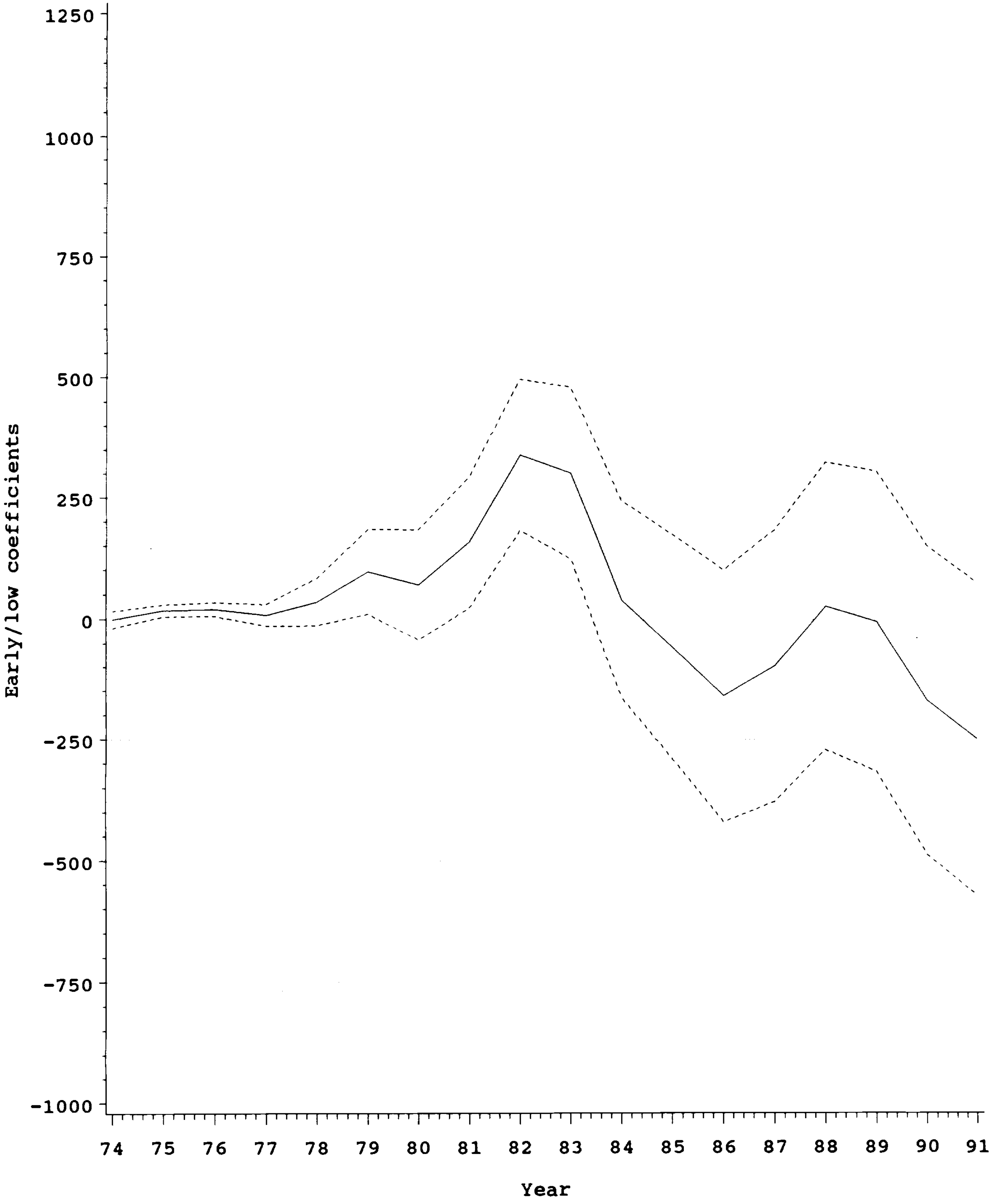

Figure 10a. Early/low-ability applicant interaction terms for Whites. Applicants in AFQT groups III and IV, who applied in 1979-82

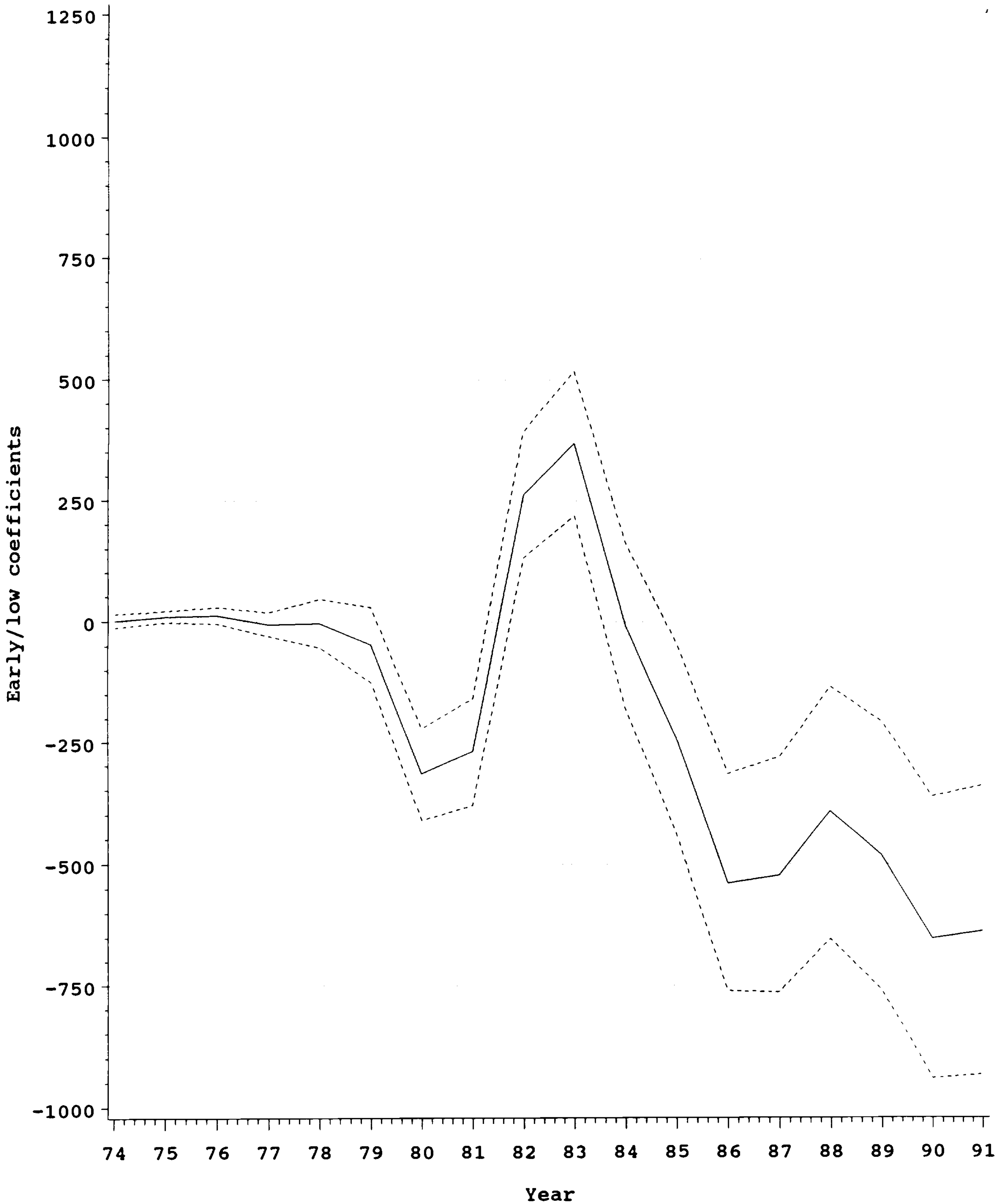

Figure 10b. Early/low-ability applicant interaction terms for Nonwhites. Applicants in AFQT groups III and IV, who applied in 1979-82

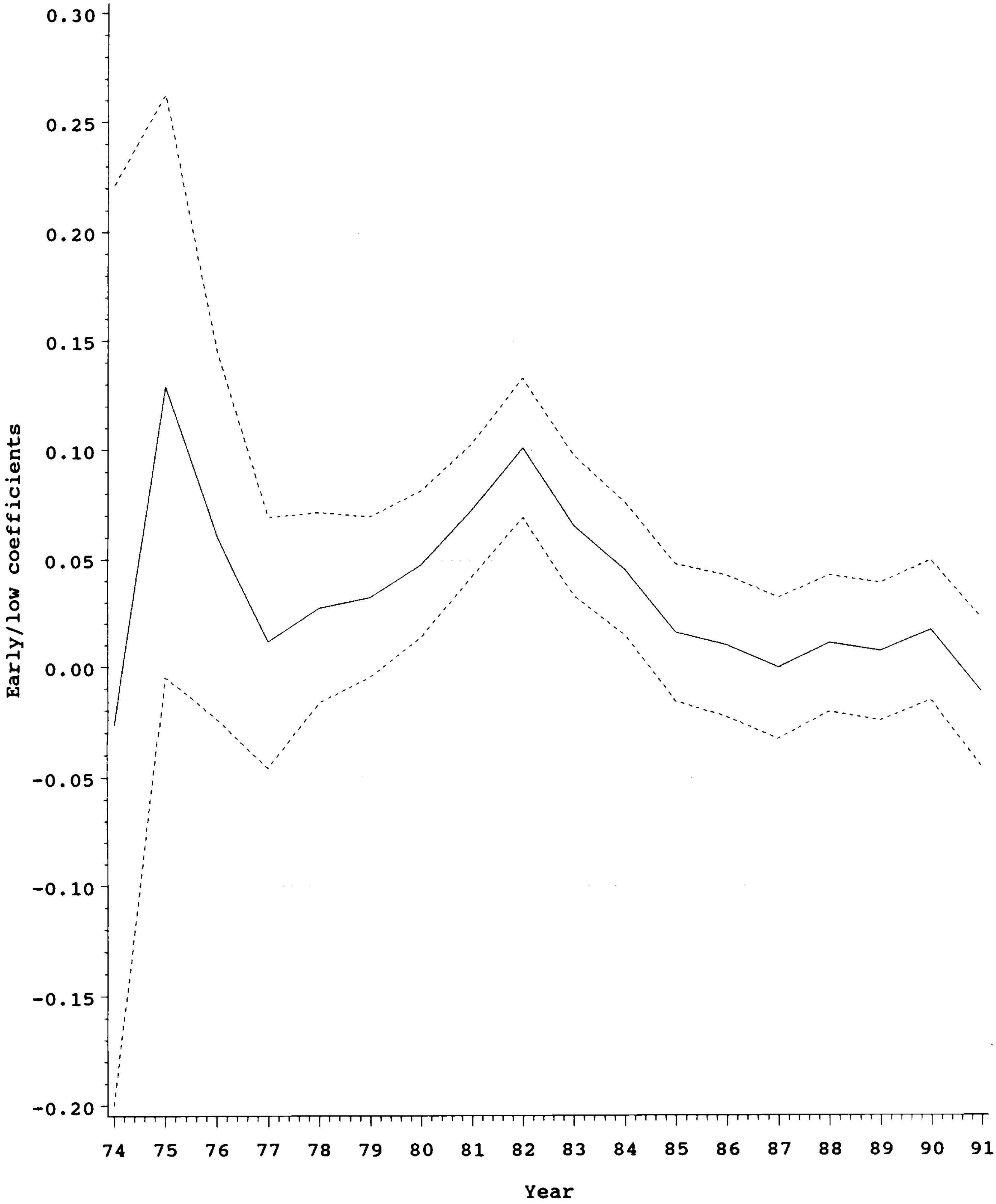

Figure 11a. LOG Early/low-ability applicant interaction terms for Whites. Applicants in AFQT groups III and IV, who applied in 1979-82

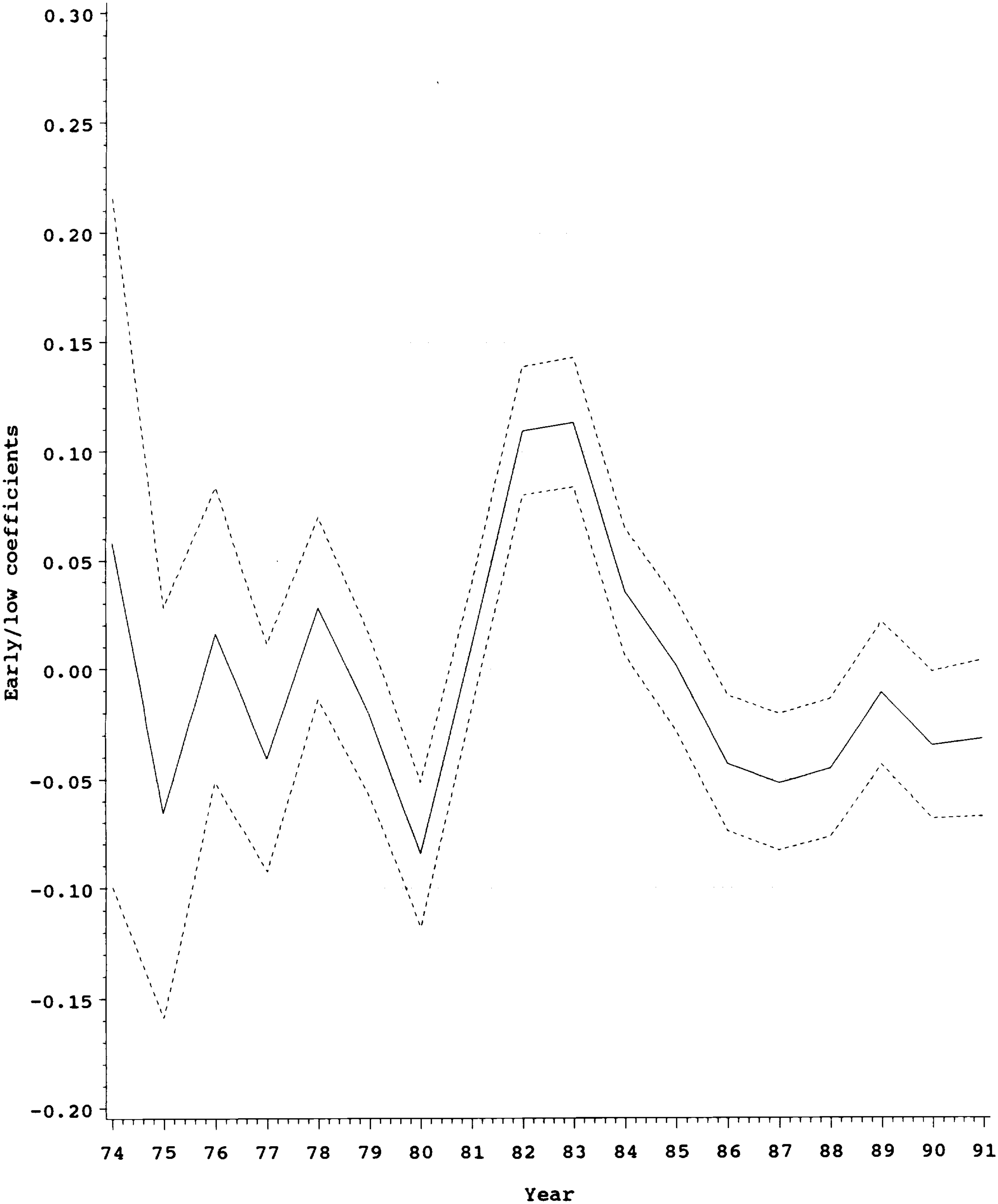

Figure 11b. LOG Early/low-ability applicant interaction terms for Nonwhites. Applicants in AFQT groups III and IV, who applied in 1979-82

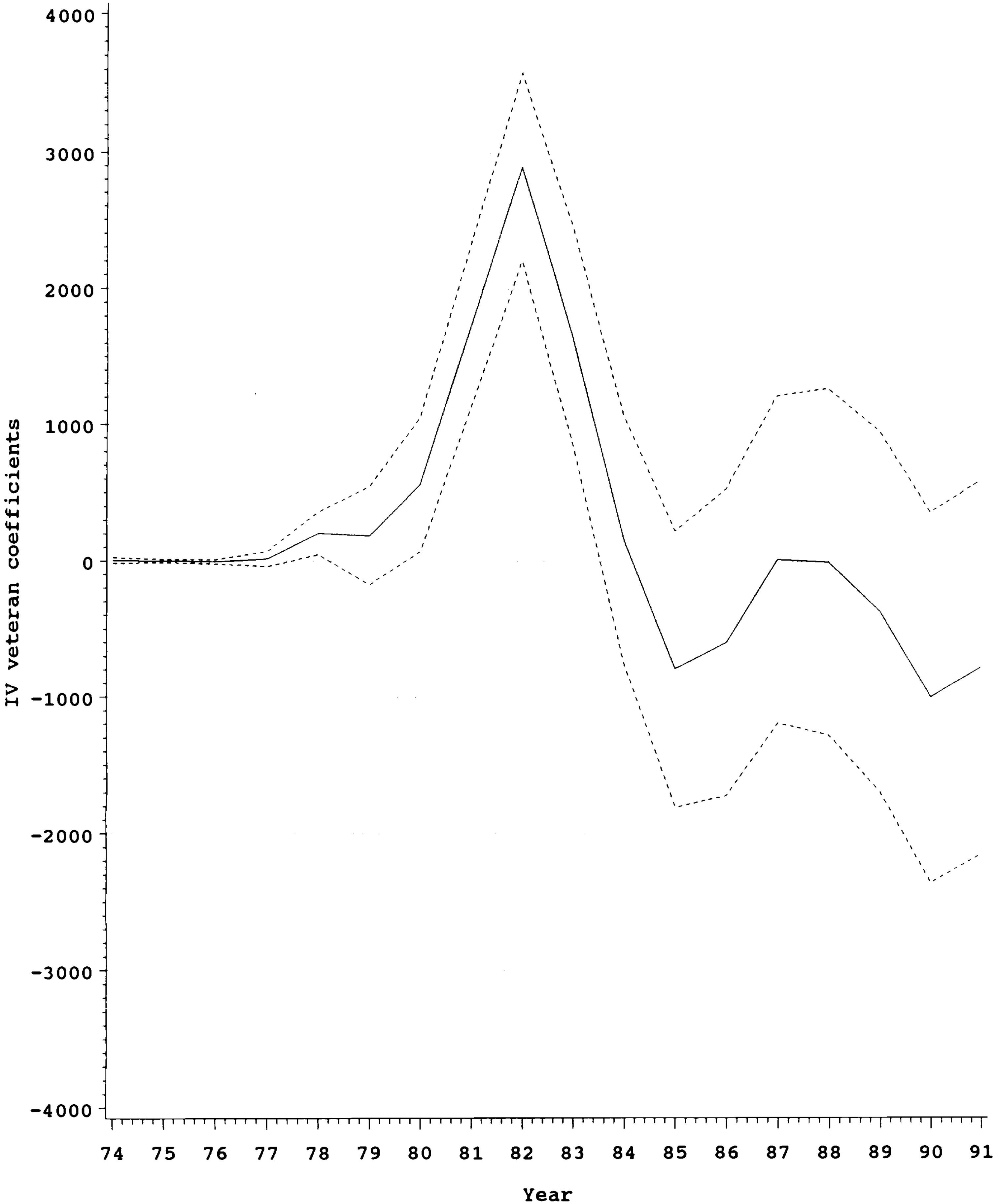

Figure 12a. Instrumental variables estimates for Whites.
Applicants in AFQT groups III and IV, who applied in 1979-82.

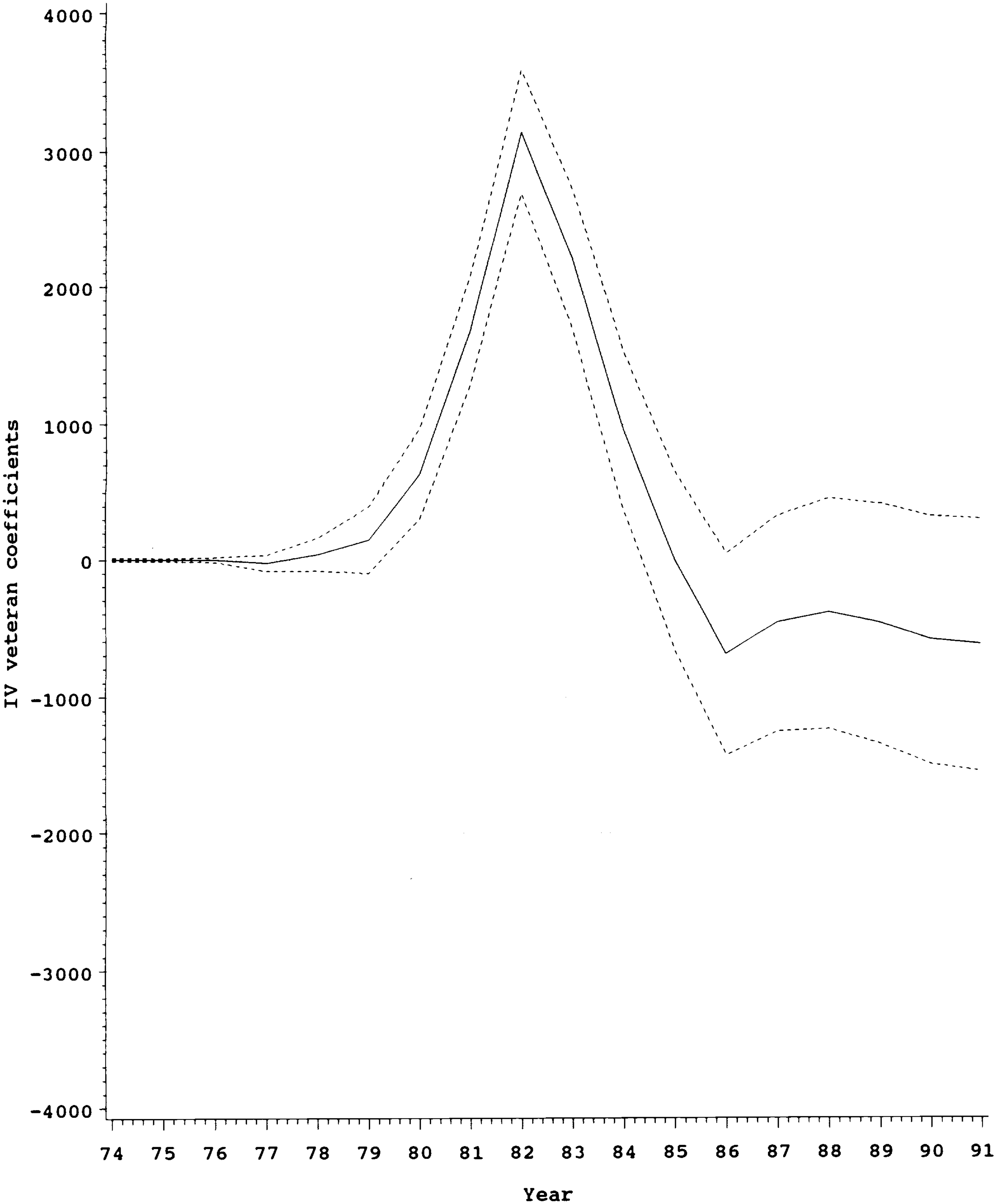

Figure 12b. Instrumental variables estimates for Nonwhites.
Applicants in AFQT groups III and IV, who applied in 1979-82.

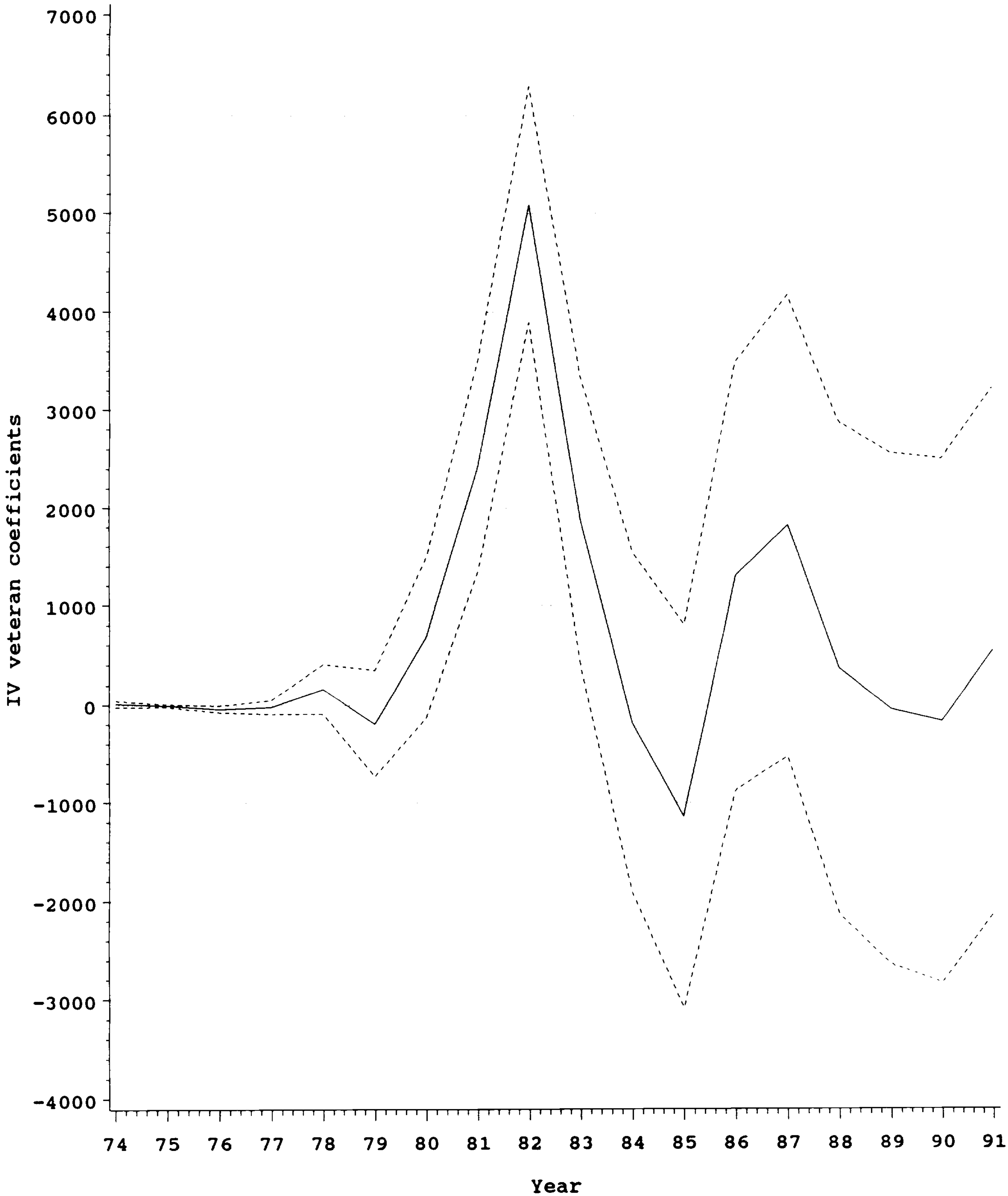

Figure 13a. Heterogeneous instrumental variables estimates for Whites. Applicants in AFQT groups III and IV, who applied in 1979-82. Saturated first stage.

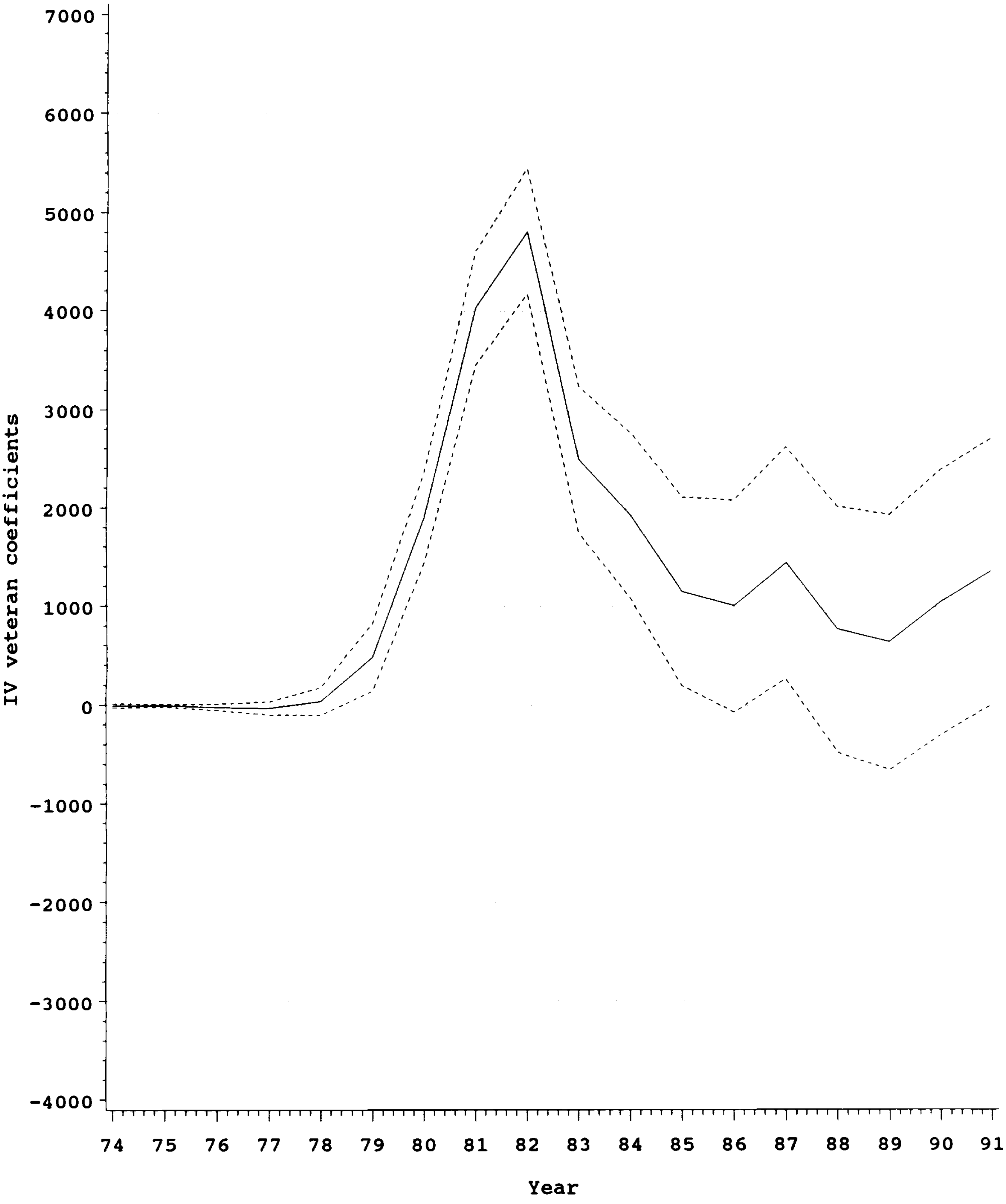

Figure 13b. Heterogeneous instrumental variables estimates for Nonwhites.
Applicants in AFQT groups III and IV, who applied in 1979-82.
Saturated first stage.

A. Statistical Appendix

I. Standard errors for $\hat{\alpha}_c$ and $\hat{\alpha}_n$.

Both $\hat{\alpha}_c$ and $\hat{\alpha}_n$ can be written in the form:

$$\text{(A.1)} \qquad \{\Sigma_k\ \delta_k N_{1k}\bar{y}_{1k}/\Sigma_k\ \delta_k N_{1k}\} - \{\Sigma_k\ \delta_k N_{mk}\bar{y}_{0k}/\Sigma_k\ \delta_k N_{mk}\},$$

where m = 0 or 1 depending on the estimator. The variance of (A.1) was computed treating δ_k, N_{1k}, and N_{mk} as fixed and assuming that veteran and nonveteran earnings are statistically independent. The variances of $\bar{y}_{1k}$ and $\bar{y}_{0k}$ are available for both earnings and log-earnings in the SSA/DMDC matched data set.

II. Standard errors for the reduced form and IV estimates

Consider first the instrumental variables estimates. A typical estimating equation is:

$$\text{(A.2)} \quad \bar{y}_j = \beta_0 + m(X_j)\beta_1 + m(A_j)\beta_2 + m(S_j)\beta_3 + [m(A_j)\otimes m(X_j)]\beta_{21} + [m(S_j)\otimes m(X_j)]\beta_{31}$$
$$+ E[\alpha_j |\ A_j, S_j, X_j;\ v_j=1]p_j + \bar{\epsilon}_j,$$

where $\bar{y}_j$ is the average earnings in cell j conditional on A_j, S_j, and X_j (but not conditional on veteran status). The data were aggregated to this level using population frequency counts. Minimum variance estimates of equations like (A.2) are the same as weighted least squares estimates using the reciprocal of the variance of $\bar{\epsilon}_j$ as weights (this is a straightforward implication of results in Newey 1990.) To simplify the variance calculation, the weights were computed here using the fact that the p_j are known population proportions. The residual variance in (A.2) is therefore also the variance of the dependent variable given the instruments. This variance was computed by aggregating grouped variances to the level of division defined by the instruments.

It is straightforward to show that asymptotic standard errors for the optimal weighted least squares estimates are given by the probability limit of the inverse weighted cross-product matrix of instruments as cell sizes approach infinity. Diagonal elements of this matrix can be recovered from conventional weighted least squares regression output by dividing the reported standard errors by the reported RMSE. Standard errors for the corresponding reduced form estimates were computed using a similar procedure. Note that since p_j is known, the same residual variance obtains for both the reduced form and structural equation estimates.

To see whether this approach provides a good measure of sampling variance, I conducted a small monte carlo experiment using the 1991 earnings of nonwhites. Earnings in each cell were resampled assuming cell averages are normally distributed around the cell mean with variance given by the variance of earnings in the cell divided by the cell size. Similarly, to replicate random draws of the weights used in weighted least squares estimation, the cell variance was also sampled from a normal distribution. The variance of the sampling distribution of cell variances was calculated assuming earnings are normally distributed, in which case the fourth moment of earnings is twice the variance squared. The cell variances were sampled independently of the cell means. The results suggest that the analytic standard errors calculated as described above provide a good approximation to the actual sampling variance of the instrumental variables estimates. Results from the same experiment suggest that confidence intervals based on the normal approximation provide reasonably accurate coverage of the population treatment effect.

Finally, note that because the first-stage equation for p_j is saturated, the instrument list includes many interaction terms. This raises the possibility of finite-sample bias because of over-fitting. In this case, however, the fact that p_j is a non-stochastic population probability means that if the model is correctly specified, estimates of equations like (A.2) are actually unbiased in finite samples.

B. Data Appendix

I. Sample design

The 753,095 records submitted to SSA to be matched to earnings were selected as a stratified random sample. First, the population of 3,023,642 applicants meeting the criteria described in the text was grouped into cells defined by the following variables:

Race = white, nonwhite	2 values
Year of application = 1976 - 1982	7 values
AFQT Group = I, II, IIIa, IIIb, IVa, IVb, IVc, V	8 values
Veteran status = veteran, nonveteran	2 values
Schooling completed at time of application = College graduate, some college, high school graduate, GED certified, grade 11, grade 9 or 10.	6 values
Year of birth = 1954 - 1965 up to 6 values for each application year (e.g. 54-59 for 1976 applicants)	
Qualification status for applicants in 1977-78 = 2 values in 1977-78 only.	

This grouping generated 8,760 non-empty population cells. To satisfy SSA confidentiality requirements, population cells with less than 25 applicants were not sampled. There are 5,654 cells with 25 or more observations in this breakdown.

To ensure that the sampling scheme did not miss small cells and that no sampled cell had fewer than 15 observations, sampling was of three types. Random samples were drawn from large cells, but all observations in cells with between 25 and 34 observations in the population were selected. Medium-sized cells were over-sampled so as to generate an expected cell size of 35. Let S be a uniformly distributed random variable and let COUNT denote the population cell size. The sampling algorithm worked as follows:

If COUNT < 25 then do not sample, otherwise
If race = white then draw S, and
    if COUNT greater than or equal to 200 then sample if S < .15,
    otherwise,
    if COUNT is less than 200 and greater than or equal to 25
    then sample if S < (35/COUNT.)
If race = nonwhite then draw S, and
    if COUNT greater than or equal to 75 then sample if S < .46,
    otherwise,
    if COUNT is less than 75 and greater than or equal to 25
    then sample if S < (35/COUNT.)

This generated a sample of 352,035 white applicants and 401,060 nonwhite applicants.

The minimum sample cell size for white applicants is 17, and the minimum sample cell size for nonwhite applicants is 15. The median cell size for white applicants is 40 and the median sample cell size for nonwhite applicants is 57. The sample is self-weighting for those observations in population cells with 200 or more observations for whites and 75 or more observations for nonwhites. Roughly 93 percent of each racial group falls in cells in the self-weighting part of the sample.

SSA programmers succeeded in locating and validating earnings data for 697,944 out of the 753,095 DMDC applicants searched for on the SER. Matches were validated using SSA data on sex, race, and year of birth. A proposed match was assumed to be valid if the DMDC and SSA sex codes matched, and if the Social Security and DMDC years of birth differed by no more than 3 in absolute value. Note that the empirical sampling rate in each cell differs from the self-weighting rates of 15 and 46 percent because of random variation, and because of over-sampling of small cells and lost observations in the earnings match.

Inflating the matched sample observations by the inverse empirical sampling frequency shows that the 697,044 records that were matched and validated using SSA data represent 848,708 nonwhite applicants and 2,149,718 white applicants. Thus, (848.71/862.6 =) 98.4 percent of the nonwhite applicants in the population are represented in the earnings match, and (2149.7/2161.1 =) 99.5 percent of the white applicants are represented.

II. Social Security earnings coverage and taxable maximum[18]

OASDI and HI (Medicare) are contributory programs in which covered workers must pay a tax and report earnings to the SSA in order to be eligible for benefits. These reports are the source of the earnings data used here. Since 1978, the SSA has received earnings reports from employers on a copy of the W-2 forms used for federal income tax purposes. Previously, employers and self-employed workers were required to file quarterly reports directly with SSA.

OASDI and HI tax rates are set for employees under the Federal Insurance Contribution Act (FICA) as a flat rate paid in equal parts by employers and employees and applied to covered earnings up to a taxable maximum. Earnings up to the taxable maximum originating in employment covered by the OASDI or HI programs are said to be FICA-taxable. Since 1991, the FICA-taxable maximum for HI has been much higher than that for OASDI. For the purposes of this paper, it is important to note that all FICA-taxable earnings, whether reported for the purposes of OASDI or HI, and whether reported as part of a mandatory or voluntary coverage provision, should appear on the SER as FICA wages. Self-employment earnings are taxed for OASDI and HI under the Self-Employment

[18]This section draws on US Department of Health and Human Service (1993.)

Contributions Act (SECA) but SECA wage data were not used in this project.

About 95 percent of all jobs in the US are currently covered by the OASDI program. Exceptions fall into five major categories: (1) Federal civilian employees hired before 1984, (2) railroad workers, (3) some employees of State and local governments already covered under a retirement system, (4) household and farm workers with low earnings, and (5) persons with no wage and salary earnings and very low earnings from self-employment. Members of the uniformed services have been covered since 1956, and receive noncontributory wage credits to improve their insured status. Recent important changes in coverage include the coverage of most federal employees and employees of nonprofit organizations in 1983, HI coverage of many state and local employees in 1986, OASDI coverage of State and local employees without an employer retirement plan in 1990, and coverage of members of reserve components of the military in 1987.

Column 1 of Table A1 shows the percent of all wage and salary earnings (excluding self-employment earnings, but including earnings above the taxable maximum) originating in FICA covered employment. This figure fluctuates around 90 percent. Column 2 reports the FICA taxable maximum for OASDI and HI, including the new higher HI maximum for 1991. Column 3 of the table shows the fraction of covered male workers with total annual earnings below the taxable maximum, which is also around 90 percent in recent years.

III. SSA confidentiality edit

The confidentiality edit accepted as input the uncensored grouped data set showing cell-identifiers and earnings variables for the 5,654 cells defined above. Cells were masked to ensure that the file released satisfied SSA confidentiality requirements. Masking was implemented according to the following algorithm for each of the 18 years of earnings data in each of the 5,654 cells in the sample:

Let EARN = average earnings
POSEARN = average positive earnings
LOGERN = average log earnings
STD = standard deviation of earnings
STDPOS = standard deviation of positive earnings
STDLOG = standard deviation of log earnings
ZERO = number in cell with zero earnings
XMAX = number in cell with earnings exactly at the taxable maximum
GEMAX = number in cell with earnings at or above the taxable maximum
FREQ = cell count (in sample)

1. If {(FREQ-ZERO)*POSERN} < 1001 then mask:
EARN, POSERN, LOGERN, STD, STDPOS, STDLOG, ZERO, GEMAX, XMAX
2. If (FREQ-ZERO) < 3 then mask:
EARN, POSERN, LOGERN, STD, STDPOS, STDLOG, ZERO
3. If ZERO < 3 then mask: ZERO, POSERN, STDPOS.
4. If {(XMAX > 2) and ((GEMAX-XMAX) < 3)} then mask: GEMAX.
5. If GEMAX < 3 then mask: GEMAX.
6. If XMAX < 3 then mask: XMAX.
7. If STD=0 then mask: EARN, STD.
8. If STDPOS=0 then mask: POSERN, LOGERN, STDPOS, STDLOG.

Rules 1 and 2 effectively eliminate all earnings information on cells with very low earners or few positive earners. Information on the number of individuals with zero earnings and the number with earnings at or above the taxable maximum is also masked in cells with few individuals in these categories (rules 3-6.) Information on positive earnings is masked when there are few individuals with zero earnings (rule 3.) Finally, in the event that everyone in a cell has the same earnings, earnings information in the cell is masked (rule 7-8.)

The impact of the confidentiality edit is primarily on earnings data for young men in the first few years of the sample period. For example, there are 1,099 cells with data on the earnings of men born 1962 or later. The confidentiality edit masks earnings data for 1974 in all but 51 of these cells. But only 91 cells in this group have 1978 earnings data masked, and only one cell is masked in 1980. After 1980, no cells are masked for this young group. Of the 1,459 cells for men born in 1957 or earlier, no cells have data on average earnings that was masked. Among the 3,096 cells with data on men born between 1958 and 1961, no earnings data is masked after 1976. Earnings for 1976 are masked in only 32 cells, and 1974 earnings are masked for 1,747 cells.

For the purposes of empirical work reported in this paper, I made the following modification to the SSA confidentiality editing process. When the ZERO variable in a cell is masked but the EARN variable for that cell is not masked (this situation occurs when there are substantial earnings in a cell but few zeros,) I assume there are no individuals with zero earnings in the cell. Means and standard deviations for the full edited data set are reported in Table A2.

Table A1: Social Security Coverage

Year	Coverage (% of all Wages)	Taxable Maximum	Percent of covered male workers with wages below taxmax
1974	88.6	13,200	76.2
1975	88.9	14,100	76.4
1976	89.7	15,300	76.3
1977	90.2	16,500	76.3
1978	90.5	17,700	75.4
1979	90.3	22,900	83.6
1980	89.6	25,900	85.5
1981	89.2	29,700	87.4
1982	89.7	32,400	88.3
1983	89.6	35,700	89.6
1984	90.5	37,800	89.4
1985	90.2	39,600	89.3
1986	90.5	42,000	89.7
1987	90.8	43,800	89.9
1988	91.6	45,000	89.4
1989	91.6	48,000	90.1
1990	91.4	51,300	91.0
1991	91.7	53,400 (HI: 125,000)	

Notes: Data derived from US Department of Health and Human Services (1993, Tables 3.B2 and 4.B2).

Table A2: Descriptive Statistics

Race	Year	Average Earnings (1)	Sample size (2)	Average positive earnings (3)	Number w/data on zeros (4)	Pct. w/zero earnings (5)	Pct. w/earnings ge taxmax (6)
Whites	74	1279.75 (2317.75)	207001	3391.30 (3307.62)	202956	.648	.
	75	1566.96 (2679.24)	252186	3850.63 (3730.75)	248160	.608	.011
	76	2349.75 (3062.18)	295829	4556.14 (3852.40)	281815	.519	.
	77	3587.16 (3580.69)	321669	5435.64 (4135.00)	288210	.402	.014
	78	5252.38 (4429.71)	330912	6488.90 (4776.79)	273214	.273	.024
	79	6701.46 (5222.13)	332861	7688.63 (5359.68)	273040	.177	.024
	80	7320.74 (5741.11)	333020	8217.41 (5739.88)	286173	.132	.011
	81	8418.70 (6467.63)	333041	9278.97 (6313.07)	293535	.100	.009
	82	9021.05 (7029.04)	333041	10081.83 (6723.13)	309492	.105	.015
	83	9993.40 (7752.81)	333041	11214.06 (7390.40)	311612	.104	.009
	84	11505.58 (8603.05)	333041	12771.47 (8144.25)	311942	.094	.014
	85	12659.15 (9427.52)	333041	14021.61 (8880.57)	314498	.095	.018
	86	13739.08 (10370.27)	333041	15329.12 (9764.32)	317642	.102	.015
	87	14748.10 (11032.28)	333041	16470.11 (10352.83)	319009	.104	.018
	88	15791.62 (11606.37)	333041	17654.55 (10862.33)	319060	.104	.021
	89	16382.11 (12046.45)	333041	18446.93 (11226.05)	321281	.112	.024
	90	16563.99 (12380.19)	333041	18869.03 (11505.89)	323525	.123	.025
	91	16144.80 (12655.28)	333041	18865.97 (11752.92)	327034	.146	.028

Table A2: Descriptive Statistics (cont.)

Race	Year	Average Earnings (1)	Sample size (2)	Average positive earnings (3)	Number w/data on zeros (4)	Pct. w/zero earnings (5)	Pct. w/earnings ge taxmax (6)
Nonwhites	74	877.61	262801	2801.49	262187	0.689	.
		(1952.59)		(3083.16)			
	75	1028.86	308016	3070.94	307615	0.666	.
		(2243.15)		(3424.99)			
	76	1625.99	340169	3784.87	336420	0.579	.
		(2679.13)		(3688.34)			
	77	2690.45	355199	4696.85	340144	0.456	0.005
		(3263.77)		(4038.98)			
	78	4116.61	362829	5736.17	335136	0.324	0.012
		(4075.10)		(4564.74)			
	79	5317.24	364696	6558.84	328170	0.225	0.034
		(4709.23)		(4995.82)			
	80	5899.34	364903	7104.08	340026	0.186	0.011
		(5134.91)		(5193.94)			
	81	6706.64	364903	7957.56	348604	0.165	0.006
		(5788.67)		(5654.58)			
	82	7139.25	364903	8685.46	353445	0.180	0.007
		(6301.93)		(6041.76)			
	83	7872.20	364903	9549.43	355816	0.175	0.011
		(6964.85)		(6660.96)			
	84	8945.46	364903	10552.39	353669	0.151	0.012
		(7633.11)		(7275.07)			
	85	9784.71	364903	11454.82	355502	0.146	0.008
		(8323.95)		(7938.76)			
	86	10459.93	364903	12340.96	356676	0.152	0.011
		(9054.25)		(8643.34)			
	87	11069.72	364903	13073.31	357327	0.154	0.011
		(9584.67)		(9150.06)			
	88	11595.36	364903	13724.48	358033	0.157	0.009
		(10093.76)		(9646.80)			
	89	11806.52	364903	14179.14	358913	0.169	0.010
		(10528.60)		(10081.39)			
	90	11887.92	364903	14578.48	360653	0.186	0.011
		(10827.72)		(10350.78)			
	91	11465.05	364903	14692.04	362912	0.220	0.010
		(11039.91)		(10607.40)			

Notes: Standard deviations in parentheses. Sample sizes show the number of observations in each year with earnings data (including zeros). Column 4 shows the number of observations for which data on the number of zeros in the cell is available. Column 3 shows the weighted average of positive earnings across all cells in which the number of zeros has not been censored and the weights are the number of uncensored positive observations in the cell. Column 6 is a percentage of column 2 for uncensored cells.