All bills currently being considered would change the types of services allowed, although in varying degrees. Currently CETA programs provide many services, including classroom training, on-the-job training, work experience, job placement assistance, and allowances for participants while being trained. The Administration's proposal would eliminate work experience and allowances for participants. The Senate bill would also eliminate work experience and current participant allowances. On the other hand, the House Committee bill would retain work experience and, for needy persons, allowances.

Previously Employed Persons with Chronically Low Earnings

None of the current types of training seemed to help persons with chronically low earnings--more often, men than women. For this group, for whom there is a smaller margin for increasing the amount of time worked, greater reliance must be placed on raising wage rates. Obtaining higher-wage jobs requires greater skills, which in turn probably requires more extensive, and thus more expensive, training. In other words, to increase substantially the future earnings of recently employed persons with chronically low earnings would require concentrating more resources on fewer individuals.

For this group, the magnitude of the potential benefits of extensive training is uncertain; however, some results from a CETA demonstration project, the Skill Training Improvement Program, that provided training for more highly skilled jobs, suggest that positive results might be possible.[10] This study of 15 prime sponsors indicated that a greater proportion of these participants obtained jobs when leaving the program, and at higher wage rates, than participants in CETA comprehensive programs. The study did not, however, examine the long-term effects and did not include a control group.

Bills currently before the Congress would allow, but not require, longer training.

---

10. See Abt Associates, Inc., STIP I: CETA and the Private Sector (prepared for U.S. Department of Labor, September 1979).

APPENDIX A. ESTIMATING THE EFFECT OF CETA TRAINING ON PARTICIPANTS' FUTURE EARNINGS

This appendix expands the discussion in Chapter III of how the effect of CETA training programs was estimated. It provides an overview of the methodology, a brief description of the data, a discussion of the statistical model, and an explanation of the estimation procedure. Further methodological issues are discussed in Appendixes B through H.

OVERVIEW OF THE METHODOLOGY

The methodology used to estimate the effect of training can be described approximately in five basic steps. The outcome of each step is presented in Table A-1.

Step 1: Estimating Participants' Post-Program Shifts From Their Past Earnings Trends

A linear trend fit through 1970-1973 annual earnings for each CETA participant was used to predict future earnings in the absence of training (see Figure A-1).[1] A $1,100 average difference (in nominal dollars) between actual and predicted post-program earnings was obtained for female participants and a $100 average difference was obtained for male participants.

Step 2: Accounting for Changing Economic Conditions

Corresponding shifts from past earnings trends were estimated for comparison group members during the post-program period. These shifts (roughly $300 for women and $200 for men, in nominal dollars) were subtracted from the results of Step 1 to account for changes in economic conditions affecting everyone during this period.

1. Figure A-1 is the same as Figure 7 in Chapter III.

TABLE A-1. ESTIMATES OF AVERAGE ANNUAL POST-PROGRAM EARNINGS GAINS AT EACH STEP IN THE ANALYSIS[a]

At the End of:	For Women	For Men
Step 1: Obtaining the Basic Finding (in nominal dollars)	1,100[b]	100[c]
Step 2: Accounting for Changing Economic Conditions (in nominal dollars)	800[b]	-100[c]
Step 3: Also Accounting for the Unusually Low Earnings in the Year Before Training (in nominal dollars)	900[b]	-100[c]
Step 4: Also Accounting for Inflation (in 1980 dollars)	1,100[b]	-100[c]
Step 5: Also Accounting for Individual Socioeconomic Differences (in 1980 dollars)	1,300[b]	200[c]

SOURCE: Estimates were derived from the Continuous Longitudinal Manpower Survey and the March 1976 Current Population Survey supplemented by individual Social Security earnings records.

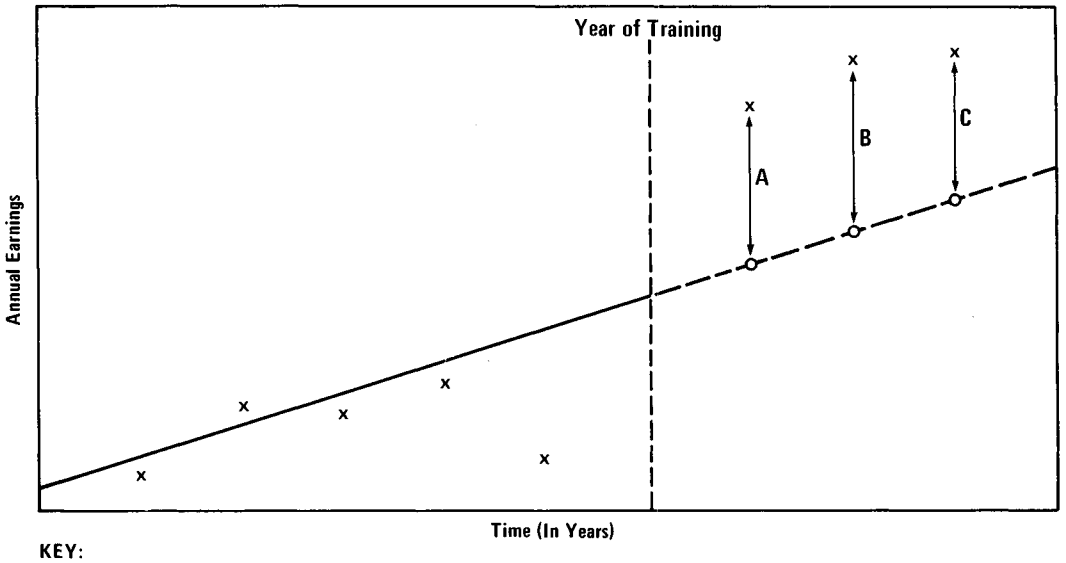
a. For persons over 24 years of age and in CETA training more than seven days.

b. Significant at the 0.01 level.

c. Not significant at the 0.05 level.

Figure A-1.

Earnings After Training Relative to the Past Long-Term Earnings Trend of a CETA Participant Who Experienced a Post-Program Earnings Gain



KEY:
x = Actual annual earnings
O = Predicted annual earnings without training
A, B, and C = Difference between actual and predicted earnings

Step 3: Accounting for Participants' Unusually Low Earnings In the Year Before They Entered Training

Male participants earned an average of roughly $1,200 (in nominal dollars) below their earnings trend, and female participants earned an average of $400 below their earnings trend in the year before they entered a training program. Appendix B estimated the rate at which these "pre-program dips" probably would have disappeared in the absence of training. This rate (which was quite fast) was used to estimate the portion of the pre-program dip that would have remained without training (less than $100, on average, for both male and female participants—see Appendix B), and the results of Step 2 were adjusted to account for this factor.

Step 4: Accounting for Inflation

Estimates of post-program earnings gains were based on data for several different years. To combine these results in consistent monetary units, they were all expressed in 1980 dollars, using the Personal Consumption Expenditures component of the implicit price deflator for Gross National Product.

Step 5: Accounting for Individual Differences in Personal Characteristics

Results to this point accounted for individual differences in past earnings trends, which in turn accounted for individual differences in measurable factors, such as age and education, and unmeasurable factors, such as motivation, that affect potential future earnings. However, to the extent that some of these measurable factors predict likely changes in behavior that produce substantial future deviations from past earnings trends, it was necessary to account for these factors explicitly in the analysis.

This was done by including a variety of personal characteristics as independent variables in the regression model described below.[2] Doing so raised the estimate of the effect of training

---

2. Personal characteristics included were age, age squared, education level, education level squared, family size, minority status, and whether or not the individual: was a household head; was currently married; was never married; had children under 4; had children between 4 and 6; and had children between 7 and 18.

for women from $1,100 to $1,300 and changed the estimate for men from an insignificant -$100 to an insignificant +$200.

THE DATA

Estimates of the effect of training were based on data from the Continuous Longitudinal Manpower Survey for 3,223 CETA participants and data from the March 1976 Current Population Survey for 30,668 comparison group members. The participant sample contained 932 women and 677 men who had been in classroom training, 236 women and 414 men who had been in on-the-job training, and 447 women and 517 men who had been in work experience programs. Tables A-2 and A-3 describe these groups.

In addition to the survey data, individual earnings data for many years before training and up to three years after training were available from each CETA participant's Social Security record. Data for two post-program years were obtained for almost all participants, and data for the third post-program year were obtained for about half of the participants. Corresponding Social Security earnings data were also obtained for all comparison group members. The validity of these longitudinal earnings data was examined in Appendix C and found to be quite satisfactory.

The sample of participants was composed of persons over 24 years old who entered CETA classroom training, on-the-job training, or work-experience programs (not public service employment) between January 1975 and June 1976 and stayed in the program more than seven days. Persons over 24 were chosen to focus on adults with meaningful past earnings experience. Participants in public service employment were excluded to focus directly on CETA's comprehensive training title. Persons entering between January 1975 and June 1976 were chosen because they were the only groups for which appropriate data were available at the time of the analysis. And persons staying in the program for more than seven days were selected both to ensure a minimal exposure to training and to be consistent with the criterion used by Westat, Inc., in previous analyses of these data.[3]

---

3. See Westat, Inc., Continuous Longitudinal Manpower Survey: The Impact of CETA on Participant Earnings, Working Paper No. 2 (U.S. Department of Labor, June 1980), p. 2-3.

TABLE A-2. CHARACTERISTICS OF FEMALE PARTICIPANTS AND COMPARISON GROUP MEMBERS IN THE SAMPLE[a]

	Classroom Training	On-the-Job Training	Work Experience	All CETA Training	Comparison Group Members
Average Number of Days in CETA	150	120	162	149	---
Average Age	34	35	37	35	41
Average Years of Education Completed	10.9	11.4	11.5	11.1	12.5
Average Number of Family Members	3.7	3.4	3.7	3.7	3.5
Percentage Minority	54	41	36	47	19
Percentage Currently Married	32	45	38	35	77
Percentage Never Married	13	12	15	13	6
Percentage Formerly Married	55	43	47	51	17
Percentage with Children Under 4	10	11	7	9	17
Percentage with Children Between 4 and 6	5	8	7	6	21
Percentage with Children Between 7 and 18	13	14	15	14	51

NOTE: All information pertains to the date of entry for CETA participants and to March 1976 for comparison group members.

a. Includes only persons who were over 24 years old and, for CETA participants, also includes only persons who were in a training program for more than seven days.

TABLE A-3. CHARACTERISTICS OF MALE PARTICIPANTS AND COMPARISON GROUP MEMBERS IN THE SAMPLE[a]

	Classroom Training	On-the-Job Training	Work Experience	All CETA Training	Comparison Group Members
Average Number of Days in CETA	132	120	139	131	---
Average Age	33	33	36	34	39
Average Years of Education Completed	11.1	11.3	10.8	11.0	12.3
Average Number of Family Members	3.5	3.2	3.4	3.4	3.2
Percentage Minority	52	32	38	42	23
Percentage Currently Married	55	62	54	57	76
Percentage Never Married	27	19	25	24	14
Percentage Formerly Married	18	19	21	19	10
Percentage with Children Under 4	8	7	6	7	21
Percentage with Children Between 4 and 6	5	4	7	6	17
Percentage with Children Between 7 and 18	13	12	12	12	36

NOTE: All information pertains to the date of entry for CETA participants and to March 1976 for comparison group members.

a. Includes only persons who were over 24 years old and, for CETA participants, also includes only persons who were in a training program for more than seven days.

Participants were divided into two groups according to when they entered the program. The first group--referred to as 1975 participants--included all persons who entered between January and August 1975. The second group--referred to as 1976 participants --included all persons who entered between September 1975 and June 1976. This split was made to minimize the timing mismatch for the definition of the year immediately before training (the year of the "pre-program dip"). Because Social Security earnings data were only reported by calendar year, 1974 was defined as the pre-program year for 1975 participants and 1975 was defined as the pre-program year for 1976 participants. The first post-program year for an individual member of either group was defined as the first full calendar year after that person left a training program.

The comparison group was defined to include all persons from the March 1976 Current Population Survey who earned less than the maximum earnings reported by Social Security records in every year between 1970 and 1975; who were between 25 and 60 years old; and who were members of families with incomes of less than $30,000 in 1975.[4]

THE STATISTICAL MODEL

The effect of CETA training was estimated from the following model:

$$Y_{it} = \alpha_i + \beta_i \cdot t + \gamma \cdot T_{it} + \sum_j \delta_j \cdot X_{ji} + \epsilon_t + \epsilon_{it} \qquad (A1)$$

and

$$\epsilon_{it} = \rho \cdot \epsilon_{it-1} + V_{it} \qquad (A2)$$

---

4. The maximum earnings covered by Social Security and thus reported by Social Security records were $7,800, $7,800, $9,000, $10,800, $13,200, and $14,100 from 1970 through 1975, respectively. Persons in families with incomes greater than $30,000 were eliminated to be consistent with analysis by Westat, Inc., who supervised development of the data base. See Westat, Inc., Continuous Longitudinal Manpower Survey: The Impact of CETA on Participant Earnings, Working Paper No. 2 (U.S. Department of Labor, June 1980), p. 2-2.

where:

Y_{it} = person i's earnings in year t;

t = time (the last two digits of calendar year t);

T_{it} = one if year t was after person i terminated from training and zero otherwise (thus it was always zero for comparison group members);

X_{ji} = the jth personal characteristic for person i;

ϵ_t = a year-specific error component reflecting economic conditions;

ϵ_{it} = the individual error component for person i in year t;

V_{it} = the random portion of person i's error component in year t;

α_i and β_i = the intercept and slope of person i's earnings trend;

γ = the average effect of CETA training on future annual earnings;

δ_j = the coefficient for the jth personal characteristic; and

ρ = a first-order serial correlation coefficient that varied by sex.

Equation A1 specifies separate earnings trends with parameters α_i and β_i for each person in the sample. In addition, personal characteristics, X_{ji}, were included to account for systematic differences in likely deviations from past earnings trends.

A dummy variable, T_{it}, was included to distinguish observations representing post-program years for participants from all other observations. Its coefficient, γ, was the average effect of training. For some of the analyses, T_{it} was replaced by a separate dummy variable to measure the effect of each major type of training.

Equation A1 specifies two error components. The first component, ϵ_t , represents the average effect of fluctuating economic conditions on everyone in the sample. The second component, ϵ_{it} , represents the result of idiosyncratic events affecting each individual.

Lastly, Equation A2 specifies a serial correlation process to represent possible relationships over time in the effects of individual idiosyncratic experiences.

The preceding model is an extension of the fixed-effect model of heterogeneous earnings functions used by Kiefer (1979).[5] The fixed-effect model specifies one person-specific parameter for each individual to account for unique characteristics that cannot be measured directly. Equation A1, however, specifies two person-specific parameters per individual to account for unmeasured factors affecting both the underlying level and the change over time in individual long-run earnings potential.

In addition, the model is a direct extension of covariance models used in econometrics to pool multiple time-series (see Pindyck and Rubinfeld (1976)) and is an application of interrupted time-series analysis used widely for evaluation research (see Campbell (1975)).[6]

THE ESTIMATION PROCEDURE

The model was estimated in several stages in order to accommodate separate intercepts and slopes for each person and to incorporate the serial correlation structure.

---

5. See Nicholas M. Kiefer, "Population Heterogeneity from Panel Data on the Effects of Vocational Education," Journal of Political Economy, vol. 87. no. 5, pt. 2 (October 1979), pp. 213-26.

6. See Robert S. Pindyck and Daniel L. Rubinfeld, Econometric Models and Economic Forecasts (McGraw-Hill, Inc., 1976), pp. 203-06. Also see Donald T. Campbell, "Reforms As Experiments," in Elmer L. Struening and Marcia Guttentag, eds., Handbook of Evaluation Research, Vol. 1 (SAGE Publications, 1975), pp. 75-86.

Estimating Individual Trends

Estimates $\hat{\alpha}_i$ and $\hat{\beta}_i$ of the intercepts and slopes, α_i and β_i, were obtained from pre-program-year earnings data for 1970 through 1973 (excluding the year immediately before participants began training). Thus by definition:

$$\hat{\alpha}_i + \hat{\beta}_i \cdot t \equiv \alpha_i + \beta_i \cdot t + W_{it} \tag{A3}$$

where W_{it} is a random error. Substituting Equation A3 into Equation A1 yields:

$$(Y_{it} - \hat{\alpha}_i - \hat{\beta}_i \cdot t) = \gamma \cdot T_{it} + \sum_j \delta_j \cdot X_{ji} + \epsilon_t + \epsilon_{it} - W_{it} \tag{A4}$$

or

$$\hat{DEV}_{it} = \gamma \cdot T_{it} + \sum_j \delta_j \cdot X_{ji} + \epsilon_t + \epsilon_{it} - W_{it} \tag{A5}$$

where $\hat{DEV}_{it}$, person i's observed deviation in year t from his or her estimated trend, can be computed directly from post-program earnings data for participants and corresponding data (for 1976 through 1978) for comparison group members.

Incorporating Serial Correlation

The next step was to incorporate the serial correlation process, in order to account for the unusually low earnings experienced by participants in the year before they entered training. This was accomplished as follows. Equation A2 implies that:

$$\epsilon_{it} = \rho^s \cdot \epsilon_{it-s} + V_{it}' \tag{A6}$$

where t-s is the year before entering a CETA program (or a corresponding year for comparison group members), t is a post-program year (or a corresponding year for comparison group members), and V_{it}' is a linear combination of person i's random individual error components for years t-s through t. Substituting Equation A6 into Equation A5 yielded:

$$\hat{DEV}_{it} = \gamma \cdot T_{it} + \sum_j \delta_j \cdot X_{ji} + \epsilon_t + \rho^s \cdot \epsilon_{it-s} + V_{it}' - W_{it} \tag{A7}$$

Because ϵ_{it-s} and T_{it} were correlated (participants had a pre-program dip but comparison group members did not), ignoring ϵ_{it-s} would produce estimates of γ, the effect of training, that

contained a bias proportional to ρ^S. Because ρ appeared to be quite small (see Appendix B) this bias was probably also quite small. Nevertheless it was eliminated as follows.

ϵ_{it-s} was estimated by $\hat{\epsilon}_{it-s}$, the observed deviation from the 1970-1973 trend, in the year before participants entered a training program and in a corresponding year for comparison group members.[7] ρ was estimated as described in Appendix B separately for men and women. Substituting $\hat{\rho}^S$ and $\hat{\epsilon}_{it-s}$ for ρ^S and ϵ_{it-s} in Equation A7 yielded:

$$\hat{DEV}_{it} = \gamma \cdot T_{it} + \sum_j \delta_j \cdot X_{ji} + \epsilon_t + \hat{\rho}^S \cdot \hat{\epsilon}_{it-s} + Z_{it} + V_{it}' - W_{it} \quad \text{(A8)}$$

where Z_{it} was the error in estimating $\rho^S \cdot \epsilon_{it-s}$ (which was correlated with $\hat{\rho}^S \cdot \hat{\epsilon}_{it-s}$, but was independent of T_{it} and the X_{ji}). Subtracting $\hat{\rho}^S \cdot \epsilon_{it-s}$ from both sides of Equation A8 and simplifying the notation yielded:

$$\hat{\hat{DEV}}_{it} = \gamma \cdot T_{it} + \sum_j \delta_j \cdot X_{ji} + \epsilon_t + U_{it} \quad \text{(A9)}$$

where $\hat{\hat{DEV}}_{it}$ was the deviation from trend in post-program year t, adjusted for the deviation from trend in pre-program year t-s and U_{it} equalled $Z_{it} + V_{it} - W_{it}$.

Accounting for Inflation

The next step was to express the dependent variable in 1980 dollars using the Personal Consumption Expenditures component of the implicit price deflator for Gross National Product. This yielded:

$$\hat{\hat{DEV}}_{it}^* = \gamma^* \cdot T_{it} + \sum_j \delta_j^* \cdot X_{ji} + \epsilon_t^* + U_{it}^* \quad \text{(A10)}$$

where the stars indicate values in 1980 dollars.

---

7. 1974 and 1975 were used as comparison group counterparts to the year immediately before training in proportion to their occurrence as the pre-program year for participants.

Estimating the Final Model

To estimate Equation A10 from a pooled sample of data for different post-program years for participants and corresponding years (between 1976 and 1978) for comparison group members, it was expressed as:

$$\hat{\hat{DEV}}_{it}^{*} = \gamma^{*} \cdot T_{it} + \sum_{j} \delta_{j}^{*} \cdot X_{ji} + \sum_{m} \epsilon_{m}^{*} \cdot YR_{m} + U_{it}^{*} \qquad (A11)$$

where the YR_m were separate dummy variables to represent 1976, 1977, and 1978.

Variations of Equation A11 were the basis for all estimates of the effect of training.

APPENDIX B. ACCOUNTING FOR THE UNUSUALLY LOW EARNINGS EXPERIENCED BY PARTICIPANTS IN THE YEAR BEFORE THEY ENTERED A CETA TRAINING PROGRAM

Figures 1 through 6 in Chapter III illustrate that the average earnings of CETA participants in the year before they entered training were noticeably below their past trend. To the extent that this "pre-program dip" was a temporary aberration, from which participants soon would have recovered without training, their recovery should not be counted as part of the effect of the program. On the other hand, to the extent that this dip would have remained without training, participants' observed recovery should be counted as part of the effect of training. Because of the magnitude of this phenomenon for male participants, its interpretation and corresponding treatment can affect estimates of program impact for this group substantially.

The best data available indicate that almost all of the pre-program dip would have disappeared in the absence of training by the first or second post-program year (in other words, two to three years after the dip occurred). Final estimates of the effect of training were adjusted to take this factor into account explicitly, but because very little of the dip was estimated to persist, this adjustment was minor (less than $100, on average, for both male and female participants).

The following sections examine alternative explanations for the pre-program dip, summarize estimates of the rate at which it probably would have disappeared in the absence of training, and outline how these estimates were obtained.

ALTERNATIVE EXPLANATIONS OF THE PRE-PROGRAM DIP

A temporary pre-program dip of some magnitude was almost certainly produced by the fact that eligibility for CETA programs was based on short-term rather than long-term labor market experience. This phenomenon is an example of a common statistical

artifact, generally referred to as regression to the mean.[1] It occurred as follows:

An individual's earnings in any given year can be expressed as the sum of three components:

- o a long-term component reflecting his or her underlying earnings potential;
- o a short-term component reflecting idiosyncratic events that uniquely affected him or her that year; and
- o a short-term component reflecting economic conditions that affected everyone that year.

The third component was accounted for explicitly in estimates of the effect of training through the use of a comparison group (see Appendix A) and was not a factor in the determination of the pre-program dip.

The long-term component--commonly referred to as permanent income--is a function of individual personal characteristics such as age, sex, race, training, education, motivation, and past experience. Because of differences in these and other related factors, some persons can generally be expected to earn more than others, on average.

In addition, everyone is subject to idiosyncratic events--such as being fired because of a fight with the boss; having unusually good or bad luck finding a new job; being in the right place at the right time for a promotion; or being laid off by a bankrupt employer--that produce good years and bad years relative to one's earnings potential.

Basing CETA eligibility only on recent experience eliminates many persons who have a relatively low earnings potential but have just experienced an unusually good year (i.e., persons with positive individual temporary earnings components). In addition, it establishes eligibility for participation for persons with a somewhat higher earnings potential who have recently experienced an unusually bad year (i.e., persons with negative individual

---

1. See Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research (Rand McNally and Company, 1966), p. 10.

temporary earnings components). Furthermore, even among persons eligible for CETA, those who have recently experienced an unusually bad year are probably more highly motivated to apply than are those who have recently experienced an unusually good year.

All of these factors work in the same direction to produce the result that CETA programs enroll a disproportionate number of persons who have recently experienced an unusually bad year. Thus, their average individual short-term component is negative and their overall average earnings are below their "normal" long-term level. This average negative short-term component could (and, according to the empirical results reported below, does) explain the pre-program dip. In future years, as random idiosyncratic events produce equal proportions of positive and negative individual temporary earnings components, overall average earnings will increase back to "normal" in the absence of training.

On the other hand, it is sometimes argued that persons who recently experienced a permanent decline in their future earnings recognize the permanence of this change and apply for CETA training. This would produce a pre-program dip that would not disappear in the absence of training.

This explanation is based primarily on the experience of displaced workers who have lost specialized well-paying, stable jobs in declining regions or industries and have little prospect of regaining their former economic status. But CETA participants are disadvantaged individuals with little previous work experience or with a history of low-paying, unstable jobs. Thus, the typical jobs held in the past by CETA participants were probably not specialized and were unlikely to be any more difficult to replace than other past jobs had been.

Therefore, although persons having just experienced an unusually bad year are probably more likely to apply for CETA training, it is unlikely that many CETA participants experienced large permanent declines from their already low unstable earnings patterns. Furthermore, given the high variability in individual earnings for this group, it is unlikely that applicants could determine whether or not the decline they had experienced was permanent or temporary.

SUMMARY OF THE EMPIRICAL EVIDENCE

Ideally, to measure the rate at which the pre-program dip would have disappeared in the absence of training, one must observe what happened to persons who were identical to the CETA participants, but who did not enter the program.[2] Unfortunately, without a true experiment based on random assignment of applicants to training and a control group, this was not possible.

Nevertheless, there were two independent sources of information from which to approximate the rate at which the pre-program dip would have disappeared without training: the past experience of participants, and the experience of comparison group members during the pre-to-post-program period. For both groups, yearly fluctuations in earnings disappeared quickly.

CETA participants (especially men) experienced large, unrelated year-to-year fluctuations in earnings before they entered the program. During this period, they both recovered rapidly from unusually bad years and failed to maintain the levels they reached during unusually good years.

Similarly, there was little relationship between the relative performance of comparison group members during the pre-program year (1974 or 1975) and the post-program years (1976-78). Persons doing unusually poorly or unusually well during the pre-program year were back on their trends by the post-program period, on average.

These findings were based on estimates of the correlation between deviations from trend in one year and subsequent deviations from trend. A strong positive correlation would indicate that such deviations disappeared slowly, whereas a weak positive correlation would indicate a quick disappearance. A negative correlation (which was unlikely and did not occur) would indicate a systematic pattern of good years followed by bad years and vice versa.

Weak positive correlations (that is, deviations that disappeared rapidly) were observed for all groups. Based on these correlations and the average pre-program dip of $1,200 for male CETA participants and $400 for female participants, Table B-1

---

2. These persons must be identical, on average, in terms of factors that affect future earnings.

indicates the portion of the pre-program dip that probably would have remained in the absence of training. The manner in which these results were obtained is explained later in this appendix.

Judged by the past ability of male CETA participants to recover from unusually bad years, virtually none of their pre-program dip would have remained in the absence of training. Judged by corresponding results for comparison group members, roughly 11 percent or $130 would have remained during the first year after training, 4 percent or $50 would have remained during the second year, and 1 percent or $10 would have remained during the third year, on average. Similar results were obtained for women, although they were based on a smaller pre-program dip and a slower rate at which the dip disappeared.

Results in Table B-1 for the comparison group were used to adjust estimates of the effect of training to account for the pre-program dip. This was done to reduce the chances of undercompensating for this factor.

DESCRIPTION OF HOW THE RESULTS WERE OBTAINED

Appendix A developed the following model to estimate the effect of training:

$$Y_{it} = \alpha_i + \beta_i \cdot t + \gamma \cdot T_{it} + \sum_j \delta_j \cdot X_{ji} + \epsilon_t + \epsilon_{it} \qquad \text{(B1)}$$

and

$$\epsilon_{it} = \rho \cdot \epsilon_{it-1} + V_{it} \qquad \text{(B2)}$$

where:

Y_{it} = person i's earnings in year t;

t = time (the last two digits of calendar year t);

T_{it} = one if year t is after person i terminated from training and zero otherwise;

X_{ji} = the jth personal characteristic for person i;

ϵ_t = a year-specific error component reflecting macroeconomic conditions;

TABLE B-1. ESTIMATED PORTION OF THE PRE-PROGRAM DIP THAT WOULD HAVE REMAINED IN THE ABSENCE OF TRAINING

	Women				Men			
	First Post-Program Year	Second Post-Program Year	Third Post-Program Year	Average For The Post-Program Period[b]	First Post-Program Year	Second Post-Program Year	Third Post-Program Year	Average For The Post-Program Period[b]
Based on the Past Experience of CETA Participants								
In percents	12	4	1	7	0	0	0	0
In dollars[a]	50	20	0	30	0	0	0	0
Based on the Experience of Comparison Group Members								
In percents	27	14	7	18	11	4	1	6
In dollars[a]	110	60	30	70	130	50	10	70

a. Rounded to the nearest 10 nominal dollars.

b. For an average post-program period of 2.5 years.

ϵ_{it} = the individual error component for person i in year t;

V_{it} = the random portion of person i's error component in year t;

α_i and β_i = the intercept and slope of person i's earnings trend;

δ_j = the coefficient for the jth personal characteristic;

γ = the average effect of CETA training on future annual earnings; and

ρ = a first-order serial correlation coefficient that varies by sex.

The first-order serial correlation coefficient ρ measures the extent to which ϵ_{it-1}, the deviation from trend in year t-1, persists in year t. Thus it provides a direct measure of the persistence of the pre-entry dip because:

$$\epsilon_{it} = \rho^s \cdot \epsilon_{it-s} + V_{it}' \quad , \qquad \text{(B3)}$$

where ϵ_{it-s} is the pre-entry dip (the deviation from trend in pre-entry year t-s), s is the number of years between the post-program year t and the pre-entry year t-s, and V_{it}' is a linear combination of person i's individual random error components between years t-s and t.

If ρ were one the pre-entry dip would be permanent, and if ρ were zero the dip would be entirely temporary--that is, it would completely disappear on average by the next year. Between these two extremes the dip persists over time at the rate ρ^s. For example, if ρ were 0.5 half of the dip would remain after one year, one-quarter or ρ^2 would remain after two years, and one-eighth or ρ^3 would remain after three years, on average.

Basic Approach to Estimating ρ

As previously indicated, ρ was estimated from the experience of comparison group members during the pre-to-post-program period. The estimation procedure used for this purpose is described below. The same procedure was also used to estimate ρ from the experience of CETA participants before they entered the program by redefining the time period involved.

Defining 1974 as year t-s and 1976 as year t for the comparison group, one could, in theory, estimate ρ from Equation B3. A second estimate could be obtained by redefining 1977 as year t and reestimating Equation B3. A third estimate could be obtained by redefining t as 1978. But Equation B3 requires data on the true deviations from trend, ϵ_{it} and ϵ_{it-s}, which are not directly measurable.

Fortunately, estimates ($\hat{\epsilon}_{it}$ and $\hat{\epsilon}_{it-s}$) of these deviations could be obtained and were used instead. These estimates were obtained by computing a linear trend (estimating α_i and β_i) from individual earnings for four years before t-s and computing deviations from this trend in years t and t-s.

$\hat{\epsilon}_{it}$ was then regressed on $\hat{\epsilon}_{it-s}$ as follows:

$$\hat{\epsilon}_{it} = \theta + \psi_s \cdot \hat{\epsilon}_{it-s} + V_{it}'' \tag{B4}$$

Intuitively it would seem that ψ_s from Equation B4 is the same as ρ^s and thus provides a direct estimate of ρ. This is not true, however. Because $\hat{\epsilon}_{it}$ and $\hat{\epsilon}_{it-s}$ were based on the same trend, which was estimated with error, they contained a common error component, which produced an artificial positive correlation between them. Thus ψ_s simultaneously reflected two correlations: the artificial correlation due to the common error component and the true serial correlation.

But estimates of ψ_s can be transformed to eliminate the artificial correlation and produce consistent estimates of ρ. These transformations are presented below.[3]

For t corresponding to the first post-program year (two years after t-s):

$$\psi_2 = \frac{2.5 - 0.1\rho - 0.8\rho^2 - 2.7\rho^3 + 0.6\rho^4 + 0.5\rho^6}{2.5 - \rho - 1.5\rho^2 - \rho^3 + \rho^4} \tag{B5}$$

For t corresponding to the second post-program year (three years after t-s):

$$\psi_3 = \frac{3.0 - 0.15\rho - 2.2\rho^2 - 1.05\rho^3 + 0.4\rho^4 - 0.5\rho^5 + 0.5\rho^7}{2.5 - \rho - 1.5\rho^2 - \rho^3 + \rho^4} \tag{B6}$$

---

3. A derivation will be made available upon request.

And for t corresponding to the third post-program year (four years after t-s):

$$\psi_4 = \frac{3.5 - 0.2\rho - 2.6\rho^2 - 2.4\rho^3 + 2.7\rho^4 - \rho^5 - 0.5\rho^6 + 0.5\rho^8}{2.5 - \rho - 1.5\rho^2 - \rho^3 + \rho^4} \quad \text{(B7)}$$

Corresponding expressions for ρ in terms of ψ_s were not readily available, so a numerical conversion table was produced (see Table B-2). Working from this table, estimates of ψ_s obtained by applying ordinary least squares to Equation B4 were transformed to obtain consistent estimates of ρ .

Findings for Women and Men

The preceding analysis was conducted for the comparison group based on the relationship between their 1976, 1977, and 1978 (year t) deviations and their 1974 (year t-s) deviations from their 1970-1973 trends. Resulting estimates of ψ_s from Equation B4 and corresponding transformed values of ρ from Equations B5, B6, and B7 are presented in Table B-3.

Estimates of ρ ranged from 0.20 to 0.43 and averaged 0.33 for male comparison group members. Thus, on average, roughly 33 percent of a deviation from trend remained after one year, 11 percent remained after two years, 4 percent remained after three years, and 1 percent remained after four years. Deviations from trend for women appeared to disappear less quickly, with estimates of ρ ranging from 0.45 to 0.57 and averaging 0.52.

A similar analysis was based on the past earnings history of CETA participants. The experiences of 1975 and 1976 participants were considered separately.[4] For 1975 participants, year t was defined as 1971, 1972, or 1973, year t-s was defined as 1969, and the trend period was 1965-1968. For 1976 participants, year t was defined as 1972, 1973, or 1974, year t-s was defined as 1970, and the trend period was 1966-1969.

Results for both groups of participants were similar, yielding serial correlation coefficients averaging 0.06 for men and 0.34 for women. These coefficients were smaller than their

4. As explained in Appendix A, 1975 participants were defined as persons who entered CETA programs between January and August, 1975. 1976 participants were defined as entrants between September 1975 and June 1976.

TABLE B-2. THE RELATIONSHIP BETWEEN ρ AND ψ_S

ρ	ψ_2	ψ_3	ψ_4
0.00	1.000	1.200	1.400
0.05	1.019	1.221	1.424
0.10	1.040	1.242	1.448
0.15	1.063	1.264	1.472
0.20	1.087	1.287	1.496
0.25	1.112	1.310	1.520
0.30	1.139	1.335	1.545
0.35	1.167	1.360	1.570
0.40	1.195	1.387	1.596
0.45	1.224	1.415	1.624
0.50	1.253	1.445	1.652
0.55	1.282	1.475	1.682
0.60	1.311	1.506	1.714
0.65	1.339	1.538	1.747
0.70	1.367	1.570	1.780
0.75	1.392	1.601	1.815
0.80	1.415	1.630	1.848
0.85	1.436	1.657	1.879
0.90	1.452	1.679	1.906
0.95	1.463	1.694	1.926
1.00	1.375	2.250	2.667

SOURCE: Computations were based on Equations B5, B6, and B7.

TABLE B-3. ESTIMATES OF ψ_s AND ρ

	Women		Men	
	ψ_s	ρ	ψ_s	ρ
Based on the Past Experience of CETA Participants				
For 1975 participants[a]				
s=2	1.1497	0.32	1.0856	0.20
s=3	1.3436	0.32	1.2226	0.05
s=4	1.4254	0.05	1.4150	0.03
For 1976 participants[b]				
s=2	1.2262	0.45	0.9987	0.00
s=3	1.3926	0.41	1.2202	0.05
s=4	1.6366	0.47	1.4125	0.03
Based on the Experience of Comparison Group Members				
s=2	1.2964	0.57	1.2106	0.43
s=3	1.4713	0.54	1.3598	0.35
s=4	1.6225	0.45	1.4976	0.20

a. Participants over 24 years old who entered between January and August, 1975, and stayed in the program more than seven days.

b. Participants over 24 years old who entered between September 1975 and June 1976, and stayed in the program more than seven days.

counterparts for comparison group members. As indicated previously, however, the average results for comparison group members (ρ = 0.3 for men and ρ = 0.5 for women) were used to adjust for the pre-program dip in estimates of the effect of training (see Appendix A). This was done to reduce the likelihood of undercompensating for this factor.

APPENDIX C. VALIDATING THE SOCIAL SECURITY EARNINGS DATA

Two potential problems with Social Security earnings data are frequently cited--changes over time in the coverage of certain occupations, and truncation due to the upper bound for reported earnings. The following analysis indicates that these factors probably had a negligible effect on estimates of the effect of CETA training reported in this paper.[1]

CHANGES IN SOCIAL SECURITY COVERAGE

Not all jobs are covered by Social Security and thus not all earnings are reported by Social Security records. Therefore, earnings reported by Social Security records could artificially change because of changes in coverage (due to the increase over time in occupations covered by Social Security or due to individual shifts from covered to uncovered employment or vice versa). If this phenomenon occurred and if it affected CETA participants and comparison group members markedly differently, it could bias estimates of the effect of CETA training programs.

The likely magnitude of this potential bias was approximated by comparing Social Security earnings data for CETA participants with survey-based earnings data for the same group during the year before and the first year after training.[2] Comparable data for comparison group members were not available because survey earnings data were available only for one year.

1. The truncation problem examined below only concerns measurement error in the dependent variable. It does not include the statistical problem of truncated samples.

2. Survey-based earnings data were obtained from the Continuous Longitudinal Manpower Survey. Survey data were also available for the second year after training, but only for a portion of the sample.

Table C-1 reports the ratio of mean earnings according to Social Security data relative to mean earnings according to survey data.[3] In almost all cases, this ratio was close to one, indicating a high degree of consistency between the two independent sources of earnings data.[4]

More important, however, is the fact that this ratio changed very little over time. It was 1.05 for female CETA participants in the year before training and 1.00 in the year after training, representing a 5 percent decrease (see Table C1). Corresponding results for male participants were 0.97 and 0.93, or a 4 percent decrease. Roughly comparable changes were experienced by participants in each of the different types of training.

To estimate the likely bias due to the preceding decreases in Social Security earnings coverage requires an estimate of the corresponding shift in coverage for comparison group members. Direct information about this shift was not available. But because there was no change in the Social Security law affecting the occupations that were covered during the analysis period, there was probably no shift in average coverage for the comparison group.

Thus the maximum likely relative shift for participants (the difference between their shift and that of comparison group members) was a 5 percent decline for women and a 4 percent decline for men. Based on the average first post-program-year earnings of $4,300 and $6,800 for female and male participants, respectively, these declines imply maximum likely negative biases of roughly $200 and $300 in estimates of the effect of training--not enough to affect the conclusions of this paper.

---

3. Earnings were expressed in constant dollars to control for the slight timing mismatch between the Social Security and survey-based earnings data.

4. Both Social Security and survey-based earnings data contain random measurement error and thus are not as consistent for each individual as they are for group averages. These individual errors cancel each other and do not affect group averages appreciably. Because estimates of the effect of training are based on group averages, they are not biased by individual random measurement error in the dependent variable.