Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs

Author(s): Orley Ashenfelter and David Card
Source: *The Review of Economics and Statistics*, Vol. 67, No. 4 (Nov., 1985), pp. 648-660
Published by: The MIT Press
Stable URL: https://www.jstor.org/stable/1924810
Accessed: 04-04-2020 20:38 UTC

USING THE LONGITUDINAL STRUCTURE OF EARNINGS TO ESTIMATE THE EFFECT OF TRAINING PROGRAMS

Orley Ashenfelter and David Card*

Abstract—We use the longitudinal structure of earnings of trainees and a comparison group to estimate the effectiveness of training for participants in the 1976 CETA programs. We fit a components-of-variance model to earnings of the comparison group and use a simple model of program participation to predict the earnings histories of the trainees. These predictions provide an estimate of the effect of training and an overidentification test of the model. Our program estimates are very sensitive to the model of participation (ranging from $200 to $2000), and we conclude that randomized clinical trials are necessary to reliably determine program effects.

PASSAGE of the Manpower Development and Training Act (MDTA) of 1962 inaugurated a new series of training programs designed to raise the earnings of unemployed and low-income workers. Ten years later, despite the absence of any clear experimental test of the effectiveness of the MDTA programs, Congress implemented the Comprehensive Employment and Training Act (CETA), to the accompaniment of broad claims that the new programs would be more effective than the old. Once again without any clear experimental evidence, Congress replaced CETA with the Job Partnership Training Act (JPTA) in 1982. If history progresses as it has during the last two decades, however, then it will not be long before the recent claims of success for the JPTA are replaced by proposals for still another government training program.

The rise and fall of successive federal training programs underscores the need for credible and continuous evaluation of these programs. Yet, apart from the results of one genuine experiment,[1] these training programs must still be analyzed by non-experimental methods, even some two decades after they were first initiated. Any evaluation must therefore bring to bear statistical methods for untangling the actual effect of these programs from other factors that would have influenced trainee earnings even if no training had taken place.

In order to make any progress a comparison group of workers must be generated to control for economy-wide movements in earnings during and after the training period. In addition, it is clear by now that participants in training programs do not represent a random sample of the eligible population. Trainees have typically experienced a decline in their earnings, both absolutely and relative to any comparison group selected, in the period immediately prior to training.[2] These declines are hardly surprising, since program operators are instructed to enroll workers who have recently faced difficulties in the labor market, and it is precisely such workers who may be most anxious to participate. Nevertheless, this peculiar aspect of trainee earnings introduces considerable ambiguity into the determination of whether observed post-training earnings increases are a result of training or merely of the way in which workers are selected into the program.

In this paper we set out some methods that utilize the longitudinal structure of earnings of trainees and a comparison group to estimate the effect of training. The basic idea is to first estimate a time-series model of earnings determination from data on the comparison group. Then, using a very simple statistical hypothesis about program participation, we generate a complete time-series of earnings predictions for the trainees. The differences between predicted and actual post-training earnings serve as a natural estimate of the training effect. By the same token, differences between predicted and actual earnings in the pre-training periods provide a built-in test of the model of earnings generation and program participation,

Received for publication September 27, 1984. Revision accepted for publication February 8, 1985.

* Princeton University.

This research was supported, in part, by SRI International. We are grateful to Terry Johnson and Richard West for helpful discussions, and to two referees for comments on an earlier draft.

[1] We are referring here to the Supported Work Program administered by the Manpower Demonstration Research Corporation that used random assignment of individuals to treatment and control groups. The results of this experimental program evaluation are summarized in Manpower Demonstration Research Corporation (1980).

[2] This was first documented by Ashenfelter (1975, 1978) for MDTA trainees from the cohort of 1964. It has also been documented by Kiefer (1979) for trainees from cohorts in the late 1960s, by Bassi (1983) for the 1976 cohort of CETA trainees, and by LaLonde (1984) for the trainees in MDRC's Supported Work experiment of the 1970s.

and a simple check on the credibility of the estimated training effect.

As we shall see, this method is no substitute for a properly designed experimental test of the effectiveness of training, but it does provide some evidence on the empirical consistency of the estimated program effects. In the absence of experimental data, there seems to be no alternative to the adoption of this or similar methods of program evaluation, since we find that small differences in model specification can lead to remarkable differences in the estimated impact of training. Hopefully, the accuracy of these methods may eventually be the subject of experimental testing.[3]

The paper begins with a discussion of the Social Security earnings histories of 1976 enrollees in the Comprehensive Employment and Training Act (CETA) training programs. Earnings histories of a comparable group of non-trainees drawn from the March 1976 Current Population Survey are presented as a benchmark against which to judge the impact of training. We go on to analyze a number of alternative estimators of program effectiveness —starting with simple estimators and proceeding to those based on more complete models of earnings generation and program participation. For the most part, our analysis is confined to male trainees over 21 years of age in the training year. In the last section of the paper, however, we give a brief summary of estimated training effects for female trainees in the same age group.

I. Earnings Determination and Program Participation

The demographic characteristics and earnings histories of adult male trainees from the 1976 cohort of CETA participants are reported in table 1. Also recorded in this table are similar data for

[3] LaLonde (1984) tests several other evaluation methods using the Supported Work Program experimental data.

TABLE 1.—DEMOGRAPHIC CHARACTERISTICS AND EARNINGS HISTORIES OF TRAINEE AND CONTROL GROUPS: ADULT MALES

	Trainees[a]	Trainees Finished in 1976[b]	Controls[c]
1. Average Age (years)	30.9	30.9	31.1
2. Education (years)	11.5	11.5	12.5
3. Percentage Married	50.1	50.5	75.0
4. Percentage White (Non-Hispanic)	60.0	58.7	84.3
Earnings in 1967 Dollars[d]			
1970	2102 (2195)	2099 (2168)	3178 (2529)
	(.19/.07)	(.18/.07)	(.13/.20)
1971	2180 (2121)	2153 (2101)	3401 (2436)
	(.17/.09)	(.17/.08)	(.11/.24)
1972	2621 (2270)	2590 (2258)	4078 (2615)
	(.13/.07)	(.13/.07)	(.09/.24)
1973	2970 (2436)	2958 (2410)	4683 (2829)
	(.11/.05)	(.12/.05)	(.08/.21)
1974	2785 (2443)	2746 (2430)	4979 (3005)
	(.13/.03)	(.13/.03)	(.08/.15)
1975	1898 (2050)	1832 (1990)	4869 (2996)
	(.19/.01)	(.19/.01)	(.10/.16)
1976	1959 (1756)	2032 (1756)	5238 (3083)
	(.10/.01)	(.07/.01)	(.10/.18)
1977	2785 (2289)	2794 (2389)	5392 (3176)
	(.12/.01)	(.13/.02)	(.10/.20)
1978	3052 (2628)	3014 (2636)	5238 (3298)
	(.17/.03)	(.17/.03)	(.13/.25)
Sample Size:	3072	2161	5238

Note: All demographic variables are recorded as of 1976.
[a] The trainee sample consists of the 1976 cohort of CETA trainees from the Continuous Longitudinal Manpower Survey whose program termination dates were in 1976 or 1977.
[b] Trainees whose program termination dates were in 1976 only.
[c] The control sample consists of a stratified random sample of eligible members of the 1976 Current Population Survey. Eligibility requirements are listed in footnote 4 of the text.
[d] For each year, the column lists the mean of earnings in 1967 dollars together with the standard deviation of earnings in parentheses and the proportion of the sample with earnings equal to zero or the maximum of Social Security earnings underneath.

adult males from a sample of the March 1976 Current Population Survey.[4] In order to control for age differences between the trainee sample and the population as a whole, we have resampled the Current Population Survey to generate a control sample with the same age distribution as the trainees. As one might expect, this age adjustment does not fully eliminate the differences between the two samples in race or marital status characteristics. Our approach below is to handle these differences by a time-series model of the earnings process that contains a separate fixed effect for each individual.

Since the earnings data for the trainees and the comparison group are drawn from Social Security records, some individuals are recorded with only partial earnings information. In addition, individuals whose earnings exceed the maximum taxable earnings level are recorded as having earnings at the maximum. For each year we report the mean and standard deviation of deflated Social Security earnings, as well as the fraction of workers who are at the taxable maximum (which varies over the years) and who report no taxable earnings whatever. Workers with no taxable earnings may be earning income outside of the Social Security tax system or may be unemployed, although there is no way to determine which of these phenomena is more important.[5] In our analysis, therefore, we have included all earnings records, with no adjustment for earnings that are equal to zero or the taxable maximum.

The trainee earnings in table 1 display the characteristic pattern of a decline in real earnings in the year immediately prior to training. Indeed, the real earnings of the trainees are $200 less in 1975 than in 1970, while the earnings of the comparison group increased by some $1700 over that period. Also as expected, the level of real earnings of the trainees is always lower than the level of real earnings of the comparison group. Moreover, the difference between the earnings of these two groups widens over the nine year period, and this widening begins several years before the onset of training.

Table 1 also contains data for the subset of trainees who finished training during 1976. The table reveals few differences between this group and the entire trainee sample. It should be clear, however, that the temporal pattern of earnings for either group of trainees in table 1 is very different from the temporal pattern for the comparison group. We turn next to the simplest models of time-series earnings and program participation that might be consistent with both trainee and comparison group earnings histories.

A. Simple Models

Suppose that earnings of the i^{th} individual in period t, y_{it}, follow a simple components-of-variance scheme:

$$y_{it} = \omega_i + d_t + D_{it}\beta + \epsilon_{it} \tag{1}$$

where ω_i is a permanent component, d_t is an economy-wide component, D_{it} is a dummy variable for participation in training during period τ, taking the value of unity for trainees in the post-training periods ($t > \tau$), β represents the effect of training, and ϵ_{it} is a serially uncorrelated transitory component of earnings. It is obvious from (1) that if assignment to training is independent of ω_i and ϵ_{it}, then a simple post-training difference in earnings between trainees and controls will estimate the training effect β.

The data in table 1 reveal that this difference in earnings is surely inadequate as an estimate of the training effect. At a minimum we must allow for the fact that the trainee and comparison groups have different permanent components of earnings.

To accommodate this fact, suppose that participation in training in period τ is governed by the magnitude of the permanent component of earnings, with

$$D_{it} = 1 \quad \text{for } t > \tau \text{ if and only if } \omega_t < \bar{y},$$

[4] Eligible members of the Current Population Survey (CPS) sample satisfy the following restrictions: (1) They had to report 1975 earnings less than $20,000, and 1975 household income less than $30,000. (2) They had to report themselves in the labor market (either with a job or unemployed and looking for a job) in March 1976. The trainee and CPS samples were provided to us by SRI International. Restrictions (1) and (2) eliminate some 21% of the overall CPS population. Details on the construction of the trainee and CPS samples are provided in Dickinson, Johnson and West (1984), pp. 37–45.

[5] Among the major groups of employees outside of the Social Security Tax System are federal workers (prior to 1982) and certain state and local workers. Since the CETA programs placed many trainees in state and local employment, it is conceivable that CETA trainees' earnings records contain a disproportionate number of zeros in the post-training period. In 1976 (the year of training) the proportion of zero earnings among trainees and controls was equal to 10%. In later years, trainees had a slightly higher incidence of zero earnings, although we have no information on their post-training employment status.

where $\bar{y}$ is a constant based on potential trainees' discount rates, time horizons, and tastes for training. In this case, a simple estimate of the training effect is obtained from a comparison of the change in earnings for the trainees between some pre-training period $(\tau - j)$ and the post-training period $(\tau + 1)$ relative to the change in earnings for the control group over the same period. This "difference-in-differences" estimator provides an unbiased estimate of the training effect because

$$E(y_{i\tau+1} - y_{i\tau-j} | D_{i\tau+1} = 1) \\ - E(y_{i\tau+1} - y_{i\tau-j}) = \beta(1 - p) \tag{2}$$

for all $j > 0$, where p is the fraction of the total population that participates in training. If p is small, as is the case for virtually all training programs, then the difference-in-differences of earnings between trainees and controls provides a straightforward estimate of the training effect.

The important point to observe about this method for estimating the training effect is that as many estimates may be calculated as there are pre-training observations on earnings. Moreover, these estimates should be similar if the model is correctly specified. Calculating all of the possible estimates and comparing their values therefore provides a test of the specification of the earnings function (1) and/or of the selection rule based on permanent components of earnings.

The first column of table 2 contains estimates of the training effect for 1978 earnings based on this simple difference-in-differences method using the years 1970–1975 as base years. It is immediately apparent that these alternative estimates of the training effect are *all* different from one another. The third column contains the estimates of the training effect for 1977 earnings using only those trainees from the 1976 cohort who had completed training prior to 1977. These estimates of the training effect also differ from one another. The variability in the estimates makes it clear that either the specification of equation (1) or the selection rule based on permanent earnings components is not capturing some important elements of the data.

It should be clear that minor changes in the selection rule still lead to the prediction that many of the estimated training effects in table 2 should be similar, so long as equation (1) is maintained. Suppose, for example, that selection is based on the rule

$$D_{i\tau+1} = 1 \quad \text{if} \quad y_{i\tau-k} \leq \bar{y}$$
$$= 0 \quad \text{if} \; y_{i\tau-k} > \bar{y}.$$

Here, selection into training is based on actual earnings in the k^{th} period prior to the advent of training. Provided, however, that the transitory error in earnings ϵ_{it} is taken to be serially uncorre-

TABLE 2.—DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE TRAINING EFFECT FOR ADULT MALE CETA PARTICIPANTS (STANDARD ERRORS IN PARENTHESES)

Basis Year	Change in Earnings from Basis Year to 1978: Trainees Relative to Controls	Change in Earnings from Basis Year to 1978: Trainees Finished in 1976 Relative to Controls	Change in Earnings from Basis Year to 1977: Trainees Finished in 1976 Relative to Controls
1975	785	813	439
	(64)	(72)	(63)
1974	8	9	−365
	(68)	(76)	(68)
1973	−473	−499	−873
	(70)	(78)	(71)
1972	−729	−736	−1110
	(71)	(79)	(72)
1971	−965	−976	−1350
	(71)	(78)	(71)
1970	−1110	−1145	−1519
	(74)	(82)	(74)
Mean Difference:	−414	−422	−796
	(63)	(70)	(64)

Note: All figures are in 1967 dollars.

lated, it remains the case that equation (2) continues to hold for $j > k$. That is, the difference-in-differences estimator is still reasonable so long as the difference is taken from a period prior to the one used by program operators or potential participants as a basis for selection into training. Again, all the training effect estimates based on the pre-training base years prior to the selection year should be similar. This selection scheme may nonetheless account for differences in the training program estimates calculated from base years near to and far from the date of entrance to training.

Table 2 indicates, however, that the simple difference-in-differences estimates vary substantially over all the base years listed in the table. The difference between the calculations based on the 1975 base year and the other base years is most dramatic, but it is clear that a simple selection bias analysis, using equation (1) and assignment to training on the basis of observed pretraining earnings, is inadequate to explain the data in table 2.

One possible explanation for the apparent variability in the training effect estimates in table 2 has been advanced by Heckman (1978) and Heckman and Robb (1982). They observe that if selection is based on earnings in period $\tau - k$, then the transitory component of trainee earnings will be abnormally low in that period. They also observe that if the transitory component ϵ_{it} is serially correlated, then trainee earnings will be abnormally low in periods adjacent to $\tau - k$, returning to their permanent level only as the transitory shock wears off. To the extent that transitory earnings components in alternative basis years are more or less correlated with the negative transitory earnings component in the selection year, difference-in-differences estimates based on different pre-training years can be expected to yield different estimates of the effect of training.

Heckman (1978) and Heckman and Robb (1982) also suggest an ingenious generalization of the simple difference-in-differences estimator to cope with the autocorrelation in the transitory component of trainee earnings. Suppose that the conditional expectation of earnings subsequent to training, $y_{i\tau+1}$, given earnings in the selection period, $y_{i\tau-k}$, is linear in the latter. Then it is easy to establish that

$$\begin{aligned} E(y_{i\tau+1} | y_{i\tau-k} < \bar{y}) &= E(y_{i\tau+1}) + b(y_{i\tau+1}, y_{i\tau-k}) \\ &\quad \cdot \{ E(y_{i\tau-k} | y_{i\tau-k} < \bar{y}) - E(y_{i\tau-k}) \}, \end{aligned} \tag{3}$$

where $b(z_2, z_1)$ indicates the population regression coefficient of z_2 on z_1. Likewise,

$$\begin{aligned} E(y_{i\tau-2k-1} | y_{i\tau-k} < \bar{y}) &= E(y_{i\tau-2k-1}) + b(y_{i\tau-2k-1}, y_{i\tau-k}) \\ &\quad \cdot \{ E(y_{i\tau-k} | y_{i\tau-k} < \bar{y}) - E(y_{i\tau-k}) \}. \end{aligned} \tag{4}$$

By choosing to calculate (4) for the same number of periods "behind" the selection period $\tau - k$ as $\tau + 1$ is "ahead" of the selection period, we can guarantee equality of the regression coefficients $b(y_{i\tau+1}, y_{i\tau-k})$ and $b(y_{i\tau-2k-1}, y_{i\tau-k})$ so long as the earnings process is covariance stationary. It follows immediately that the symmetric difference-in-differences

$$\begin{aligned} &E(y_{i\tau+1} - y_{i\tau-2k-1} | D_{i\tau+1} = 1) \\ &\quad - E(y_{i\tau+1} - y_{i\tau-2k-1}) = \beta(1 - p) \end{aligned}$$

is a straightforward estimator of the training effect that handles the autocorrelation in the transitory component of earnings.

Before it is possible to implement this procedure, however, it is necessary to decide which period to take for $\tau - k$; that is, which period's earnings to use as the basis for selection into training. One logical possibility is to use $k = 1$ and assume that the selection is based on earnings in the period immediately preceding the training period. This is the information that will certainly be available to the potential participants and to the program operators. Alternatively, we may consider taking $k = 0$ and using the period of training as the selection period. Although earnings in the training period are never fully realized, the worker or program operator may have information on several months of data from which an excellent forecast may be made.

For the full cohort of trainees, only one symmetric difference-in-differences estimate of the effect of training is available, based on the difference between 1978 earnings and 1974 earnings if 1976 is taken as the selection year, or based on the difference between 1978 earnings and 1972 earnings if 1975 is taken as the selection year. As can be seen from table 2, however, these two estimates of the effectiveness of training are dramatically different, ranging from nearly zero to a statistically significant $-\$736$.

For CETA trainees whose program termination dates were in 1976, two symmetric difference estimates of the training effect are available for each selection year; one based on 1977 earnings, and one based on 1978 earnings. If the true training effect were the same in the two years then a simple specification test for the symmetric difference estimator would be to compare the training effect estimates for these two years, as they should be similar. Taking 1976 earnings as the basis for selection into training, the two symmetric difference-in-differences estimates of the training effect are \$9 and \$439, based on earnings growth from 1974 to 1978 and from 1975 to 1977, respectively. Using 1975 earnings as the basis for selection into training, on the other hand, the two estimates are −\$736, based on earnings growth from 1972 to 1978, and −\$873, based on earnings growth from 1973 to 1977. Neither pair of estimates is identical, although the estimates using 1975 as the selection year are closer together. Again, the estimates are positive when 1976 earnings are taken as the basis for selection into training, and negative and statistically significant when 1975 is used as the selection year.

In our opinion, simple difference-in-differences techniques give unconvincing estimates of the value of training for adult male CETA participants. On the one hand, while a convenient specification test of the simple (nonsymmetric) difference-in-differences estimator is available from the long span of pre-training data, the underlying assumptions for this estimator are clearly violated.[6] On the other hand, in the absence of several years of post-training data, no similar specification check is available for the symmetric difference-in-differences estimator. It is clear that arbitrary and largely unverifiable maintained hypotheses are necessary to select a symmetric difference estimator, and that different maintained hypotheses lead to very different conclusions on the value of training.

One way to provide for a test of specification is to focus more explicitly on the considerable amount of additional data available in the period prior to training for both the trainee and the comparison groups. The symmetric difference-in-differences estimator makes very little use of this information. Our approach is to use equation (3), but to recognize explicitly that the regression coefficient in this expression will vary systematically as different comparisons are made. Given a particular assumption about the structure of the earnings equation (1), the regression coefficient in (3) may be calculated explicitly from data on the comparison group alone. Since most conventional components-of-variance models of earnings contain very few parameters, this model will be highly over-identified and readily susceptible to specification tests. In effect, we will continue to use equation (3) to adjust the earnings of trainees for sample selection, but we will discipline the process by which the adjustment factor is obtained by requiring its consistency with a components-of-variance explanation for the comparison group's earnings and the pre-training earnings of the program participants.

[6] It is worth noting that the variability in estimated program effects clearly observed in table 2 was not observed in Ashenfelter's (1978) study of the 1964 cohort of MDTA trainees. Apparently the earnings structure and/or the selection mechanism for trainees has changed so much that the evaluation task is considerably more difficult with the later group.

B. Components of Variance and Selection Bias

We begin by setting out a simple model of earnings determination and program participation. Suppose, as before, that earnings are described by an additive components-of-variance scheme, with a person-specific fixed effect, a year effect, and a person- and year-specific transitory earnings component ϵ_{it}. Suppose also that ϵ_{it} is first-order autoregressive with variance σ_ϵ^2 and first-order autocorrelation coefficient α. Finally, assume that training occurs during period τ if and only if

$$y_{i\tau-k} + v_i < \bar{y},$$

where $\bar{y}$ is a constant and v_i is a random variable, assumed to be independent of any earnings components. Substituting for $y_{i\tau-k}$ from equation (1), training occurs if and only if

$$\begin{aligned} z_i = (\omega_i - \omega) + \epsilon_{i\tau-k} + v_i \\ < \bar{y} - \omega - d_{\tau-k} \equiv z, \end{aligned} \tag{5}$$

where ω represents the mean of the permanent earnings component ω_i.

Our procedure is to use the earnings function (1) and the selection rule (5) to describe the means and covariances of the time series of earnings for both program participants and controls. These predicted moments are directly comparable to the observed moments of the data, and method of moments estimation techniques can be used to obtain estimates of the parameters of the earnings process, including the training effect associated with program participation.[7]

Assuming that the control sample is approximately a random sample of the population as a whole, the means and covariances of controls' earnings are described by the unconditional moments:[8]

$$E[y_{it}] = \omega + d_t,$$
$$\text{var}[y_{it}] = \sigma_\omega^2 + \sigma_\epsilon^2,$$
$$\text{cov}[y_{is}, y_{it}] = \sigma_\omega^2 + \alpha^{|t-s|}\sigma_\epsilon^2,$$

where σ_ω^2 represents the cross-sectional variance in the permanent earnings component ω_i.

For the participant sample, on the other hand, the means and covariances of earnings correspond to conditional moments, given that selection criterion (5) is satisfied. Following Heckman and Robb (1982), we assume that the conditional expectation of participant earnings in any period is a linear function of the selection variable z_i. (This will be the case, for instance, if ω_1, ϵ_{it} and v_i are jointly normally distributed.) It follows that

$$E[y_{it}|z_i < z] = E[y_{it}] + \frac{\text{cov}[y_{it}, z_i]}{\text{var}[z_i]} \times E[z_i|z_i < z],$$

and therefore

$$E[y_{it}|z_i < z] = E[y_{it}] + D_{it}\beta + \{\text{cov}[\omega_i, z_i] + \text{cov}[\epsilon_{it}, z_i]\}\lambda^*,$$

where $\lambda^* = -E[z_i|z_i < z]/\text{var}[z_i] > 0$. The mean of trainee earnings differs from the mean of control earnings by a training effect plus the sum of two components, each of which is proportional to the number λ^*. These two components reflect the covariance of the selection variable with the underlying components of earnings. Using the definition of the selection variable z_i we can calculate these covariances and obtain the following expression for the mean of trainee earnings in period t:

$$E[y_{it}|z_i < z] = E[y_{it}] + D_{it}\beta - [\sigma_\omega^2 + \alpha^{|t-\tau+k|}\sigma_\epsilon^2]\lambda^*. \quad (6)$$

In pre- and post-training periods, the discrepancy between trainee and control earnings consists of a permanent component, $\sigma_\omega^2\lambda^*$, and a geometrically declining component, centered around the selection period, $\alpha^{|t-\tau+k|}\sigma_\epsilon^2\lambda^*$. The relative magnitude of these two selection bias components, however, is completely determined by the parameters of the earnings process, and can be estimated directly from information on the controls' earnings. The model imposes the restriction that in both pre- and post-training periods, earnings of the trainees and of the controls diverge in a systematic pattern with only one free parameter: the number λ^*.

The first column of table 3 presents the results of fitting the simple components-of-variance scheme represented by equation (1) to the means and covariances of control earnings from 1970 to 1978. The estimation method minimizes a quadratic form in the deviations of the actual from the fitted moments, with the deviations weighted by the inverse matrix of third and fourth moments of the data. On the basis of the earnings data for the control sample, much of the observed cross-sectional variation in earnings represents the effect of transitory shocks. The estimated cross-sectional variance of the permanent component ω_i is less than half the estimated variance of the transitory earnings component. The estimate of α is around 0.8, however, implying that transitory earnings shocks are quite persistent.

Once we have estimated the parameters of the earnings process and selected the period $\tau - k$ to be used for the selection year, it is a straightforward matter to calculate an estimate of trainee earnings in any period using equation (6). The only unknown parameter is the selection bias parameter λ^*, which must be inferred from a comparison with actual trainee earnings. The estimated differences between trainee and control earnings, based on the parameters from the first column of table 3, are presented in the first two columns of table 4 for the case of selection into training on the basis of 1976 ($k = 0$) and 1975 ($k = 1$) earnings, respectively. To assist in the interpretation of these predicted differences we have arbitrarily scaled the number λ^* so that 1975 earnings are predicted exactly. It should be clear,

[7] See in particular Chamberlain (1982) on the application of method of moments estimation to panel data. This strategy for joint estimation of the earnings process and selection equation was proposed by Abowd (1983).

[8] Formally, in post-training periods the means and covariances of earnings for a random sample of the population include a weighted training effect. We assume that the proportion of the population that participated in training is negligible.

TABLE 3.—MINIMUM DISTANCE ESTIMATES OF THE EARNINGS PROCESS AND TRAINING EFFECT: ADULT MALE CETA PARTICIPANTS (STANDARD ERRORS IN PARENTHESES)

	Fitted to Controls Only[a]		Fitted to Controls and Trainees[b] (selection based on 1975 earnings)		Fitted to Controls and Trainees[b] (selection based on 1976 earnings)	
	No Trend Component (1)	Trend Component (2)	No Trend Component (3)	Trend Component (4)	No Trend Component (5)	Trend Component (6)
1. Variance Components[c]						
a) Permanent Component	195.5	141.7	142.6	157.7	91.6	135.4
	(10.7)	(17.0)	(14.6)	(19.3)	(14.7)	(20.3)
b) Transitory Component	460.9	445.7	499.9	425.5	549.7	446.0
	(11.4)	(15.4)	(14.5)	(17.4)	(14.9)	(18.5)
c) Trend Component	—	3.81	—	4.46	—	4.36
		(0.37)		(0.40)		(0.41)
d) Covariance of Permanent and Trend Components	—	12.17	—	9.81	—	10.37
		(1.50)		(1.66)		(1.69)
2. Autoregressive Parameter of Transitory Earnings	0.78	0.75	0.78	0.73	0.80	0.75
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
3. Selection Bias Parameter	—	—	5.28	3.87	6.35	4.27
			(0.09)	(0.09)	(0.12)	(0.10)
4. Training Effect	—	—	−1160	41	−343	747
			(56)	(65)	(56)	(68)
5. Function Value[d]	2269	1059	2410	1114	2444	1139
(degrees of freedom)	(42)	(40)	(47)	(45)	(47)	(45)

Note: All figures are in 1967 dollars.
[a] The model is fit jointly to the means and covariances of control group earnings.
[b] The model is fit jointly to the means and covariances of control group earnings and the means of participant earnings The covariances of trainee earnings are not fit
[c] For notational convenience, variances and covariances are scaled in 10,000's of 1967 dollars.
[d] Under the null hypothesis of a correct model, the optimized function value is asymptotically distributed as a chi-square variate with degrees of freedom indicated in parentheses.

TABLE 4.—ESTIMATED RELATIVE TRAINEE EARNINGS BASED ON THE STRUCTURE OF CONTROL GROUP EARNINGS: ADULT MALE CETA PARTICIPANTS

	Predicted Difference in Earnings: No Trend Component of Variance[a]		Predicted Difference in Earnings: with Trend Component of Variance[b]		
	Selection Based on 1975 Earnings (1)	Selection Based on 1976 Earnings (2)	Selection Based on 1975 Earnings (3)	Selection Based on 1976 Earnings (4)	Actual Difference in Earnings (5)
1970	1426	1600	1149	1214	1076
1971	1588	1744	1395	1466	1221
1972	1804	1935	1685	1755	1457
1973	2089	2189	2030	2091	1713
1974	2468	2525	2450	2489	2194
1975	2971	2971	2971	2971	2971
1976	2468	3563	2681	3563	3279
1977	2089	2971	2491	3258	2607
1978	1804	2525	2377	3063	2186

Note: All figures are in 1967 dollars.
[a] Predicted trainee earnings based on estimated components of variance in column (1) of table 3. For convenience, the selection bias component is scaled to predict 1975 earnings difference exactly.
[b] Predicted trainee earnings based on estimated components of variance in column (2) of table 3. For convenience, the selection bias component is scaled to predict 1975 earnings difference exactly.

however, that alternative methods of scaling λ^* lead to essentially the same qualitative conclusions.

The structure of the model implies that the permanent component of earnings accounts for a fixed difference between the earnings of trainees and the comparison group of \$928 or \$1158, depending on whether 1975 or 1976 is used as the selection year. The transitory component is symmetric around the selection year and is considerably larger than the permanent component of the predicted earnings difference around the period of training. The strong persistence in the transitory component of earnings implies that the predicted transitory component of the earnings difference will eventually decay, but that it lasts many years. The implicit training effect estimate in table 4 is nothing more than the shortfall of the predicted control/trainee earnings difference from the actual control/trainee earnings difference in column (5) of the table. For 1978 this is \$339, if selection is based on training period (1976) earnings, and $-$\$382, if selection is based on pre-training period (1975) earnings.

The specification of this simple model may be examined by comparing the predicted and actual comparison group/trainee earnings differences prior to training. These should, of course, be similar. As can be seen from table 4, the predicted and actual earnings differences are dissimilar in 1974, and they increasingly diverge as we move back in time. The predicted differences are somewhat closer to the actual differences when selection is based on pre-training (1975) earnings than when selection is based on 1976 earnings.

The problem with the predictions in table 4 appears to be that they fail to capture a systematically weaker trend in the trainees' earnings than exists in the comparison group's earnings even prior to training. This suggests the possibility that the components-of-variance model (1) should be augmented to include a person-specific growth rate of earnings g_i, which is distributed across the population with mean g and variance σ_g^2. In this case

$$y_{it} = \omega_i + d_t + g_i t + D_{it}\beta + \epsilon_{it}, \tag{7}$$

with ϵ_{it} taken to be first-order autoregressive as before. The same methods may be used to estimate trainee earnings as before, but now the covariance of earnings in any year with the selection variable will depend on the time period *and* the number of periods from the selection year for which earnings are being predicted.

There are two additional findings that suggest the usefulness of the random growth component in (7). First, the dissimilarity between the symmetric difference-in-differences estimators in table 2 suggests the empirical possibility that the extent of selection bias in pre- and post-training earnings may be unequal, even between symmetric years around the selection period. This prediction is consistent with the hypothesis that mean earnings of the trainees and of the controls are permanently diverging. Second, an examination of the variances and covariances of earnings for the comparison group indicates increasing dispersion in earnings over time. This is consistent with cross-sectional dispersion in individual-specific growth rates in earnings, and inconsistent with the simple components-of-variance scheme given by equation (1).

Assuming that earnings are generated by equation (7), and selection into training is based on a combination of earnings in period $\tau - k$ plus a random selection error, training occurs if and only if

$$z_i = (\omega_i - \omega) + (g_i - g)(\tau - k) + \epsilon_{i\tau-k} + v_i < z. \tag{8}$$

Under this selection criterion, trainees will be those for whom permanent earnings are low, transitory earnings are low, and the accumulated growth in earnings is low. Trainee earnings will therefore differ from the comparison group's earnings because of a permanent component, a symmetric transitory component, and a trend component. Specifically, the expectation of trainee earnings in period t is given by

$$\begin{aligned} E[y_{it}|z_i < z] &= E[y_{it}] + D_{it}\beta - \Big[\big(\sigma_\omega^2 + (\tau - k)\sigma_{\omega g}\big) \\ &\quad + t\big(\sigma_{\omega g} + (\tau - k)\sigma_g^2\big) + \alpha^{|t-\tau+k|}\sigma_\epsilon^2\Big]\lambda^*, \end{aligned} \tag{9}$$

where, as before, $\lambda^* = -E[z_i|z_i < z]/\mathrm{var}[z_i]$ represents the ratio of the truncated mean of the selection variable to its variance. In this expression we have accounted for both the cross-sectional variance in earnings growth (σ_g^2) and any covariance between individual-specific growth rates and individual-specific permanent earnings components ($\sigma_{\omega g}$). If, for example, earnings growth is approximately proportional, then this covariance will be large and positive. On the other hand, if more rapid earnings growth is associated with lower permanent earnings, then this covariance will be negative.

As before, the variance components, σ_ω^2, σ_g^2, $\sigma_{\omega g}$ and σ_ϵ^2, and the autoregressive parameter, α, are all identified by the structure of control group earnings. In particular, the variances and covariances of control group earnings are given by

$$\mathrm{var}[y_{it}] = \sigma_\omega^2 + 2t\sigma_{\omega g} + t^2\sigma_g^2 + \sigma_\epsilon^2$$

and

$$\mathrm{cov}[y_{it}, y_{is}] = \sigma_\omega^2 + (s+t)\sigma_{\omega g} + st\sigma_g^2 + \sigma_\epsilon^2.$$

Therefore, given the parameters of control group earnings, the predicted earnings differentials between trainees and controls depend solely on the number λ^*. The selection bias model yields a simple one parameter description of the means of trainee earnings, given the means and covariances of control earnings.

Column (2) of table 3 contains the results of fitting equation (7) to the means and covariances of control group earnings. The cross-sectional variance of the individual-specific trend in earnings (normalizing to $t = 0$ in 1970) is very precisely estimated, as is the cross-sectional covariance of the permanent and trend components of earnings.[9] The addition of random trend components of earnings greatly improves the fit of the model to the control group earnings, as the goodness-of-fit statistics in the bottom row of the table indicate. This better fit reflects mainly the ability of the growth components to explain the increasing cross-sectional dispersion in control group earnings observed in the data.

Columns (3) and (4) of table 4 contain estimates of the predicted control/trainee earnings difference using the parameter estimates from column (2) of table 3, and assuming that selection into training is based on either pre-training period (1975) or training period (1976) earnings. Again, in each case we have scaled the selection bias parameter λ^* in equation (9) so as to predict the 1975 gap in earnings exactly, given the estimates of the variance components for the controls. The addition of a growth component considerably changes the interpretation of trainee/control earnings differences and the estimated training effect. In particular, a large share of the post-training gap in earnings is now attributed to the permanently lower growth rate of earnings for the trainees, and the implied training effect is correspondingly larger than when growth components are ignored. The addition of a random growth component also improves the fit of the model to the pre-training earnings. Not only does the addition of a random growth component improve the fit of the model to the comparison group, as our results in table 3 confirm, but it is also true that it improves the fit of the predicted trainee-comparison group earnings gap. It seems reasonable, therefore, to prefer the estimates based on the components-of-variance model that contains a growth effect.

The issue remains, however, of whether 1975 or 1976 is the more appropriate selection year on which to base the estimates. A comparison of columns (3) and (4) of table 4 indicates that applying the same components-of-variance model with two different selection rules leads to estimated training effects of \$191 and \$877. It is natural to inquire whether the goodness-of-fit of one of these models justifies greater confidence in its estimated training effect. A comparison indicates that the pre-training fit to the data in column (3) is better, but the difference involved is very small. In our view these data are simply not sufficient to distinguish between selection rules based on 1976 or 1975 ($k = 0$, $k = 1$) earnings.

Up to this point we have estimated the components-of-variance model on the control sample and then estimated the training effect and the selection bias parameter λ^* using the gap between trainee and control earnings. Columns (3)–(6) of table 3 contain estimates of the components-of-variance model of earnings that pool the data on the trainee and comparison groups. In columns (3) and (4) we have modelled selection into training on the basis of 1975 earnings. In columns (5) and (6), we model selection as based on 1976 earnings. In each case we have reported the parameter estimates for equation (7) fitted to the means and covariances of control group earnings, and the means of trainee earnings, with and without the addition of random growth components.

It should be made clear that the models fitted in table 3 represent an extraordinarily economical parameterization of the means and variances of

[9] The implied correlation between the trend and permanent components of earnings is 0.52.

control group earnings and the mean earnings of the trainee group. It is perhaps not surprising then that these restrictions do considerable violence to the data in a statistical sense, as reflected by the very large chi-squared statistics associated with the restrictions. In our view, however, these models do a reasonably good job of predicting the mean earnings of the trainees prior to training, and also the covariances of the comparison group. The difficulty that remains is the considerable variability in the estimated training effects associated with different model specifications.

These difficulties are highlighted by the different estimated training effects in the fourth row of the table. On one hand, assuming selection into training on the basis of 1975 earnings and ignoring random growth components in earnings, the estimated training effect is -1160 in 1967 dollars. On the other hand, assuming selection into training on the basis of 1976 earnings, and allowing for random growth components in earnings, the estimated training effect is $747. While we have fairly strong evidence from the control group to suggest the importance of random growth components in earnings, there is no such basis to choose between 1975 ($k = 1$) and 1976 ($k = 0$) as the selection year. The chi-squared statistics are somewhat more favorable for 1975, as is the informal evidence from the two-step procedures in table 4. In view of the remarkable difference between the estimated training effects, however, further research is clearly required to distinguish confidently between the estimates.

Finally, we also estimated the components-of-variance model of earnings represented by equation (7) on the means and covariances of control group earnings and the means and covariances of trainee earnings. Our parameterization of the covariance matrix of trainee earnings is based explicitly on the hypothesis of joint normality of the random variables ω_i, g_i, ϵ_{it}, and v_i. Under that maintained assumption, the formula for the (truncated) covariance of earnings in period t and period s is given by

$$\begin{aligned}\operatorname{cov}(y_{it}, y_{is} | z_i < z) &= \operatorname{cov}(y_{it}, y_{is}) \\ &\quad + \frac{\operatorname{cov}(y_{it}, z_i)\operatorname{cov}(y_{is}, z_i)}{\operatorname{var}(z_i)} \cdot v^* \qquad (10)\end{aligned}$$

where $v^* = (\nu - 1)/\operatorname{var}[z_i]$, and ν is the variance of a standard normal variate, truncated at $z(\operatorname{var}[z_i])^{-1/2}$. Given that $\nu \leq 1$, the predicted covariances of trainee earnings are less than the corresponding covariances of control earnings, since earnings in each period are positively correlated with the selection variable z_i (which is just a linear combination of earnings in period $y_{i\tau-k}$ and the random variable v_i). Comparing these estimates with the corresponding estimates that do not restrict the trainee covariances, the training effects and the estimated components of variance are generally similar. In a qualitative sense, the model represented by equation (10) appears to fit the covariances of trainee earnings rather well, although again the formal chi-squared statistics are unfavorable. The only major difference between the training effects summarized in table 3, with unrestricted trainee covariances, and those with restricted trainee covariances, concerns the relative fit of the 1975 and 1976 selection models. Fitting only the means of trainee earnings, the selection model based on 1975 earnings fits better. Fitting both means and covariances, however, the selection model based on the 1976 earnings fits better. This fact reinforces our hesitancy in choosing between the estimates.

C. *Estimates for Females*

Table 5 summarizes our estimated training effects for adult females in the 1976 cohort of CETA participants. These estimates are based on fitting equations (7) and (9) simultaneously to the means and covariances of control group earnings and the means of trainee earnings. The general pattern of the parameter estimates for males and females is very similar. The share of variance attributed to permanent earnings components is generally lower for females, however. For both groups, the estimated covariance of permanent and trend components of earnings is large, and for the females, in fact, the implied correlation coefficient between permanent and trend components ($\sigma_{\omega g}/\sigma_{\omega}\sigma_g$) is greater than one in three out of four cases. This inconsistency illustrates the difficulty of obtaining a parsimonious model of earnings that nonetheless captures the non-stationarity evident in the data. The estimated training effects for females display a similar pattern to the estimated effects for adult males. The lowest program estimates are associated with the assumptions that

TABLE 5.—SUMMARY OF ESTIMATED TRAINING EFFECTS: ADULT FEMALE CETA PARTICIPANTS (STANDARD ERRORS IN PARENTHESES)

	Selection Based on Earnings in	
	1975	1976
1. Training effect, allowing trend component of earnings with corresponding goodness-of-fit statistic	353 (47) $F = 598$	713 (49) $F = 597$
2. Training effect, allowing no trend component of earnings with corresponding goodness-of-fit statistic	298 (46) $F = 1349$	645 (47) $F = 1339$

Note: All figures are in 1967 dollars. The training effects are estimated jointly with a components-of-variance model for the means and covariances of control group earnings and the means of trainee earnings The value reported for the goodness-of-fit statistic (F) is asymptotically distributed as χ^2 with 45 degrees of freedom for the models in row 1 or 47 degrees of freedom for the models in row 2.

selection is based on 1975 earnings and that there are no individual-specific trends in earnings. The highest estimates are associated with the assumptions that participation in training is based on 1976 earnings, and that average growth rates of earnings differ between the trainee and comparison groups. An important distinction between the program estimates for males and females, however, is the wider dispersion in the male estimates across methods. Estimates for females, by comparison, are uniformly positive and lie in the interval between $300 and $700 per year (in 1967 dollars). Perhaps the greater dispersion in estimates for the males reflects the larger magnitude of the apparent selection bias in male trainee earnings and the correspondingly greater ambiguities in reconciling trainee earnings with comparison group earnings.[10] As it happens, the estimated training effects for females are not as sensitive to the inclusion or exclusion of individual-specific trend components of earnings as the estimates for males. The differences between estimated training effects using 1975 or 1976 as the basis for selection into training are still significant for females, although the goodness-of-fit statistics for the alternative choices are very nearly identical. The overall fit of either model to the female earnings data is considerably better than the corresponding fit to the male data.

II. Concluding Remarks

Despite two decades of experience with large-scale governmentally-funded training programs, properly designed experimental tests for the effectiveness of training are virtually nonexistent.[11] The sensitivity of the nonexperimental results in this paper leads us to conclude that for the evaluation of training programs experimental tests using random assignment are especially desirable. Nevertheless, since most programs must be evaluated by nonexperimental techniques, in this paper we have set out some new methods of program evaluation that generate testable restrictions on the nonexperimental data. In the absence of such restrictions it is unclear how one can distinguish among diverse estimates from alternative, and equally plausible, specifications. At an empirical level we find that different models lead to very different estimates of training effects. This underscores our belief that, in the absence of experimental data, it is important to test alternative specifications of the earnings and selection model.

For the simple selection rule/components-of-variance models we have applied to the 1976 cohort of CETA trainees, two factors appear to have a critical influence on the size of the estimated train-

[10] Bassi (1984) reaches a similar conclusion. Her analysis of the 1976 cohort of CETA trainees by sex and race indicates that selection bias and associated ambiguities in program evaluation are most pronounced for white males.

[11] The only exceptions of which we are aware are the Supported Work Program and the Denver and Louisville Work Incentive Demonstrations, administered by Manpower Demonstration Research Corporation, and the Seattle and Denver Income Maintenance Experiment Counselling and Education Subsidy programs.

ing effects. One is our assumption about the timing of the decision to participate in training, and the other is our assumption about the presence or absence of selection bias in the trend component of earnings. It seems clear that the highest priority for future research is to find a way to test whether models using different specifications for these factors can be distinguished empirically in the data. We have provided some formal and informal tests of alternative model specifications, but it appears that additional tests of model specification will be necessary for a confident assessment of the magnitude of training effects.

The informal evidence we have presented suggests that CETA participant earnings contain permanent, transitory, *and* trend-like components of selection bias. The informal evidence, however, simply does not allow us to discriminate effectively between assumptions about the year of selection into training. Formal testing, moreover, gives contradictory evidence on the appropriate assumption about the selection year. If earnings in the year prior to training are the appropriate selection criterion, then our findings suggest that the training effect for adult males who participated in CETA in 1976 is small: at most on the order of 300 current dollars per year. If earnings in the training period are the appropriate selection criterion, then the training effect is surely larger. For adult females, on the other hand, the effect of program participation is unambiguously positive, and on the order of 800–1500 current dollars per year. Further computational experience with the models used here would no doubt be valuable for testing the sensitivity of these conclusions to alternative model specifications.

REFERENCES

Abowd, John M., "Program Evaluation," unpublished manuscript, University of Chicago Graduate School of Business, 1983.

Ashenfelter, Orley, "The Effect of Manpower Training on Earnings: Preliminary Results," in *Proceedings of the Twenty-Seventh Annual Winter Meeting of the Industrial Relations Research Association* (1975), 252–260.

______, "Estimating the Effect of Training Programs on Earnings," this REVIEW 60 (Feb. 1978), 47–57.

Bassi, Laurie J., "The Effect of CETA on the Postprogram Earnings of Participants," *Journal of Human Resources* (Fall 1983), 539–556.

______, "Estimating the Effect of Training Programs with Non-Random Selection," this REVIEW 66 (Feb. 1984), 36–43.

Chamberlain, Gary, "Panel Data," National Bureau of Economic Research Working Paper Number 913, June 1982.

Dickinson, Katherine, Terry Johnson, and Richard West, "An Analysis of the Impact of CETA Programs on Participants' Earnings," Final Report Prepared for the U.S. Department of Labor Employment and Training Administration, SRI International, Mar. 1984.

Heckman, James J., "The Longitudinal Analysis of Earnings," unpublished manuscript presented at SSRC Conference on Longitudinal Studies of the Labor Market, Oct. 1978.

Heckman, James J., and Richard Robb, "The Longitudinal Analysis of Earnings," unpublished manuscript, University of Chicago, June 1982.

Kiefer, Nicholas M., "The Economic Benefits from Four Government Training Programs," in Farrell E. Bloch (ed.), *Evaluating Manpower Training Programs* (Greenwich, CT: JAI Press, 1979), 159–187.

LaLonde, Robert J., "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," Princeton University, Industrial Relations Section Working Paper No. 183 (Nov. 1984).

Manpower Demonstration Research Corporation, "The Final Report: Summary and Findings of the National Supported Work Demonstration," Manpower Demonstration Research Corporation, 1980.