NBER WORKING PAPERS SERIES

ESTIMATING THE EFFECT OF TRAINING ON EMPLOYMENT AND UNEMPLOYMENT DURATIONS: EVIDENCE FROM EXPERIMENTAL DATA

John C. Ham

Robert J. LaLonde

Working Paper No. 3912

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1991

This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3912
November 1991

ESTIMATING THE EFFECT OF TRAINING ON EMPLOYMENT AND UNEMPLOYMENT DURATIONS: EVIDENCE FROM EXPERIMENTAL DATA

ABSTRACT

Using data from a social experiment, we estimate the impact of training on the duration of employment and unemployment spells for AFDC recipients. Although an experimental design eliminates the need to construct a comparison group for this analysis, simple comparisons between the average durations or the transition rates of treatments' and controls' employment and unemployment spells lead to biased estimates of the effects of training. We present and implement several econometric approaches that demonstrate the importance of and correct for these biases. For the training program studied in the paper, we find that it raised employment rates because employment durations increased. In contrast, training did not lead to shorter unemployment spells.

John C. Ham
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260
and
Institute for Policy Analysis
University of Toronto

Robert J. LaLonde
Graduate School of Business
University of Chicago
Chicago, IL 60638
and NBER

1. INTRODUCTION

Employment and training programs often improve the labor market prospects of economically disadvantaged women.[1] This improvement results largely from increases in post-program employment rates rather than from increases in wages or in weekly hours for those who work. Training raises employment rates because former trainees find jobs faster when unemployed or hold on to their jobs longer when employed.

This paper develops an econometric framework for estimating the effect of training on the duration of unemployment and employment. There are several reasons for separately estimating these effects. First, considerable benefits may accrue from combining a program that improves the employment prospects of unemployed workers with a program that enhances the ability of employed workers to retain their jobs. Second, a program that lengthens trainees' employment durations may be preferred to one that shortens unemployment durations because stable employment is likely to lead to greater human capital accumulation than frequent job hopping. Such human capital increases may lead to subsequent rises in trainees' wages. Finally, for evaluations using short sampling frames, separate estimates of training's impact on employment and unemployment durations may be used to estimate the program's long-run effect.[2]

---

[1]Among nonexperimental evaluations Ashenfelter (1978) reports earnings gains for the 1964 MDTA cohorts; Bassi (1983) and Ashenfelter and Card (1985) find more modest earnings gains for CETA participants. Barnow (1986) provides a summary of evaluations of CETA. Card and Sullivan (1987) report gains in employment rates for CETA participants. Several experimental evaluations report employment and earnings gains for women. Those studies include: The National Supported Work Demonstration (see Hollister, Kemper, and Maynard (1984)); the WIN Research Laboratory Project (see Wolthagen and Goldman (1983)); The Work-Welfare Experiments (see Friedlander (1988)); and the AFDC Homemaker Home Health Aid Demonstration (Bell and Orr (1987)).

[2]An estimate of the long-run effect of training on employment rates can be calculated from the difference between the ratio of the expected duration of employment to the sum of the expected durations of employment and unemployment for the trainees and the corresponding ratio for the control group.

In order to examine the effect of training on the durations of unemployment and employment, we apply our econometric approach to data from a social experiment. The advantage of these data is that among the population of eligible program volunteers, a woman's training status is uncorrelated with unobserved heterogeneity. In this case, a simple comparison between trainees' and controls' employment rates yields an unbiased estimate of training's effect on short-run employment rates. But, as we shall show, similar comparisons between the durations of trainees' and controls' employment and unemployment spells yield potentially biased and economically misleading estimates of the training effect. Consequently, even when using experimental data, evaluations of training's effect on employment and unemployment durations require a parametric statistical framework.

Our empirical findings may be summarized as follows. First, in our data, the standard empirical practice of using only fresh spells to avoid initial conditions problems in event history analysis yields misleading estimates of the training effect because it creates a serious sample selection problem that contaminates the experimental design. Second, our econometric framework successfully addresses these sample selection problems. This finding is important because there is no modification to the sample design that would eliminate such problems in our data. Finally, the social experiment studied in the paper — the National Supported Work (NSW) Demonstration — raised the trainees' employment rates because it helped those who found jobs hold onto them longer. The program had no effect on the rate at which trainees left unemployment.

The remainder of the paper proceeds as follows. Section 2 describes the social experiment. Section 3 discusses the problems that occur when using experimental data to make inferences about the effect of training on employment and unemployment durations. Section 4 constructs a statistical model that formally addresses the problems raised in the previous section. Section 5 reports our empirical findings. Section 6 concludes the paper and briefly discusses the benefits of having experimental data when analyzing this type of problem.

2. TRAINING'S EFFECT ON EMPLOYMENT RATES

The NSW Demonstration provided work experience to a random sample of eligible AFDC women who volunteered for training.[3] These women were guaranteed 9-18 months of subsidized employment in jobs in which productivity standards were raised gradually over time. Most jobs were in clerical or services occupations and paid slightly below the prevailing wage in the participants' labor markets. When their subsidized employment ended, the trainees were expected to enter the labor market and find regular jobs.

Despite similar pre-program employment rates, the trainees' post-program employment rates substantially exceeded those of the control group members. As shown by Figure 1, the trainees' and controls' pre-program employment rates were essentially identical and were declining in the 48 semi-monthly periods prior to the baseline. After

[3]There are 275 trainees and 266 controls in our sample. Table A presents the means of the trainees' and controls' demographic and pre-baseline employment characteristics using the entire experimental sample. As expected with an experimental design, these means for the two groups are nearly identical for every characteristic (except marital status). For more details on the NSW sample see the appendix. For an in-depth discussion of the program and its costs see Hollister et al. (1984).

the baseline, the employment rates of the two experimental groups diverged as the trainees entered Supported Work jobs. The employment rates of the two groups approached each other as the trainees' terms in Supported Work ended or they voluntarily dropped out of the program. Nevertheless, in the 100th semi-monthly period, or more than a year after the typical trainee had left Supported Work, the employment rates of the trainees exceeded those of the control group by 9 percentage points. Thus the experimental evaluation shows that at least in the short run, NSW substantially improved the employment prospects of AFDC participants.

Supported Work achieved these employment gains by helping trainees to hold on to their jobs longer and/or to find jobs faster, thereby increasing the length of their employment spells and/or reducing the length of their unemployment spells. A seemingly natural way to analyze the separate effects of training on employment and unemployment durations consists of comparing the mean durations of the trainees' and controls' employment and unemployment spells. These comparisons indicate that the controls have longer employment and unemployment spells. Unfortunately, this finding simply reflects the longer sampling frame for the controls. (Since the trainees spend an average of one year in training, their sampling frame is approximately one year shorter than the controls' sampling frame.) Thus simple comparisons between trainees' and controls' mean durations are uninformative.

Comparisons between trainees' and controls' empirical survivor functions (i.e., the probability of remaining employed or unemployed through a given date) avoid the problem associated with comparisons between mean durations. As shown by Table 1, the trainees' employment spells are longer than the controls' spells. For example, 65 percent of the trainees' employment spells lasted six months, compared with only 57 percent of the controls' spells. By contrast, the two groups' unemployment spells are of similar lengths; for example, 73 percent of both the trainees' and the controls' spells lasted at least 6 months. These comparisons between trainees' and controls' empirical survivor functions suggest that the NSW program raised employment rates by increasing the duration of employment spells rather than by decreasing the length of unemployment spells. Unfortunately, such a simple analysis may be inappropriate. The following section discusses the problems associated with these comparisons and presents the case for adopting a more formal statistical framework when using experimental data to estimate the effect of training on employment and unemployment durations.

3. EXPERIMENTAL DATA AND DURATION ANALYSES

There are at least three reasons why, even in experimental settings, comparisons between trainees' and controls' empirical survivor functions may yield misleading estimates of the training effect. The first problem arises because the first year of data for the controls comes from a period when the treatment group is in training. Thus, on average, the controls may face different demand conditions than the treatments.

The second problem with comparing survivor functions is that there are other differences among individuals besides their training status. In general, failing to account for such differences biases our measurement of the training effect even though a woman's training status is determined by random assignment.[4] This bias occurs since the empirical survivor functions in Table 1 are a function of the corresponding hazard functions and neglected heterogeneity will bias those calculations. Therefore, even when using experimental data, heterogeneity must be taken into account when analyzing the effects of training on the duration of employment and unemployment.

The third problem with the comparison of the trainees' and controls' survivor functions occurs because much of the data on the controls' unemployment spells come from their unemployment spells which are in progress at the baseline. Data from interrupted spells are comparable to data from fresh spells only in the absence of duration dependence, and in our empirical work we find significant evidence of duration dependence.[5]

Standard empirical analysis of duration models can avoid the first two problems, first, by conditioning on demand variables and observed characteristics and second, by allowing for unobserved heterogeneity which is uncorrelated with observed characteristics, including training status. Standard empirical analysis deals with the third problem of interrupted spells by discarding them and using only fresh spells that begin after the baseline. As shown by column 5 in Table 1, if we adopt this procedure, training actually appears to have raised unemployment durations.

---

[4]See, for example, Heckman and Singer (1984a) and Lancaster and Nickell (1980). Ridder and Verbakel (1984) discusses this problem explicitly in an experimental setting.

[5]In the absence of unobserved heterogeneity, one can avoid this problem by measuring duration from the beginning of the spell in progress at the baseline and not from the baseline. However, in the presence of unobserved heterogeneity, this adjustment will still produce biased results. See Heckman and Singer (1984a). We find significant evidence of unobserved heterogeneity below.

However, by following standard practice in our study of economically disadvantaged women, we create a potentially serious sample selection problem. For a control to have a fresh unemployment spell, she must first complete the unemployment spell in progress at the baseline, and then complete an employment spell before the end of the sample period (Figure 2a). Nearly half the controls in our sampling frame never leave the unemployment spell in progress at the baseline, and thus they never appear in our sample of fresh unemployment spells. If these controls are less skilled than the controls who have fresh unemployment spells, using only fresh spells contaminates the experimental design by comparing above-average members of the control group with typical trainees.[6] As shown by Table 2, controls with fresh unemployment spells (in column 5) have more prior work experience and education than either the full sample of controls or the sample of trainees with unemployment spells.[7] Consequently, it is quite possible that the controls with fresh unemployment spells leave unemployment more quickly than the trainees because they are a select sample and not because training increases unemployment duration.

A similar sample selection problem may arise for the trainees. Although most trainees become unemployed when they leave Supported Work, some move directly into a regular job. Further, some of these trainees never experience a subsequent spell of unemployment during the sample period. As shown in columns 1 and 3 of Table 2,

---

[6]This point simply restates the general result that using only new spells is inappropriate in the presence of unobserved heterogeneity (Heckman and Singer 1984a).

[7]Two measures of work experience are reported in Table 2. The first measure is the number of semimonthly periods of employment in the two years prior to the baseline. The second measure is the fraction of women who had never had a regular job.

trainees with employment spells are more skilled than those with unemployment spells. Consequently, the sample of trainees with unemployment spells may exclude women with "above average" characteristics.

Standard duration models can account for the differences in observed characteristics resulting from the foregoing sample selection problem. However, such models cannot account for differences in unobserved characteristics, since they assume that the heterogeneity is uncorrelated with observed characteristics. By contrast, in our problem unobserved heterogeneity will be correlated with a woman's training status. Thus, to follow standard empirical practice, we would have to adopt one of the following two implausible assumptions: (i) there is no duration dependence in unemployment spells, or (ii) unobserved heterogeneity is uncorrelated with training status in the sample experiencing fresh unemployment spells. We now turn to an estimation approach that avoids both of these assumptions.

4. ECONOMETRIC MODEL

In order to develop a likelihood function that addresses the problems discussed in the previous section, we segment the data into three parts: (i) the employment and unemployment spells that began after the baseline (fresh spells); (ii) the controls' unemployment spells that are in progress at the baseline (interrupted spells); and (iii) the treatments' time in training (training spells).

4.1 The Contribution of the Fresh Spells

The contribution of a fresh spell to the likelihood function is straightforward.[8] We define the hazard functions for exiting employment and unemployment to be as follows:

$$\lambda_j(t_j | \cdot, \theta_j) = \lambda_j(t_j | X_j, D, \theta_j; \beta_j), \tag{1}$$

where $j = e$ denotes an employment spell and $j = u$ denotes an unemployment spell. In (1), t_j is the duration of the current spell, X_j is a vector of explanatory variables (some of which vary with time), D denotes a dummy variable which equals 1 when a woman belongs to the treatment group, θ_j is a scalar random variable representing unobserved characteristics, and β_j is a vector of parameters to be estimated. The foregoing hazard function gives the probability conditional on a woman's training status and observed and unobserved characteristics that she leaves employment (unemployment) in period t, given that she has been in employment (unemployment) up to that point. This probability depends of course on the corresponding density and distribution functions for the number of periods spent in an employment (or unemployment) spell:

$$\lambda_j(\cdot) = \frac{f_j(\cdot)}{1 - F_j(\cdot)} = \frac{f_j(\cdot)}{S_j(\cdot)}. \tag{2}$$

In (2), $f_j(\cdot)$ and $F_j(\cdot)$ are the density and distribution functions associated with the length of employment (or unemployment) spells, and $S_j(\cdot)$ is the survivor function. The fresh spells' contribution to the likelihood function follows from these hazard functions. For example, suppose a woman had a fresh employment spell lasting t_e periods followed by an

[8]See, e.g., Flinn and Heckman (1982) or Lancaster (1990).

unemployment spell (of t_u periods) which was in progress when the sampling frame ended (see Figure 2a). Conditional on the unobserved heterogeneity, the contribution to the likelihood for these spells is given by

$$(3) \qquad L_f(t_e,t_u|\cdot,\theta_e,\theta_u) = f_e(t_e|D,\tilde{X}_e,\theta_e;\beta_e)\cdot S_u(t_u|D,\tilde{X}_u,\theta_u;\beta_u).$$

In (3), $\tilde{X}_e$ and $\tilde{X}_u$ contain the explanatory variables for the fresh employment and unemployment spells, respectively.

4.2 The Contribution of Interrupted Spells

The interrupted spells' contribution to the likelihood is more complicated than the contribution of the fresh spells. It might seem that the two years of pre-baseline data, noted in Figure 1, would alleviate the difficulties associated with these interrupted spells.[9] Using such data, however, simply moves the initial conditions problem back to the beginning of the pre-baseline data. Further, a likelihood function that describes the pre-baseline data must account for the program's eligibility criteria. The NSW administrators required participants to be unemployed when they volunteered for training and to have been unemployed for at least three of the six months prior to the baseline. Unfortunately, the likelihood function that accounts for this eligibility criterion and these data are extremely complicated (see Appendix B). Therefore we relegate its estimation to future work.

---

[9]Use of the pre-baseline data would also increase sample size and therefore increase the precision of the estimates.

Because of the problems surrounding the pre-baseline data, we use only the post-baseline data when accounting for the interrupted spells' contribution to the likelihood. Even in this case, the exact likelihood remains extremely complicated (see Appendix B). To see how these complications arise, let M denote the event that a woman is eligible to participate in training. Consider a woman who leaves her interrupted spell after r periods and then experiences fresh employment and unemployment spells, denoted by $\bar{t}$ (see Figure 2a). Her contribution to the likelihood is given by

$$(4) \qquad L(r,\bar{t}\,|\,\cdot,\theta_e,\theta_u) = \int_{\theta} h(r\,|\,\cdot,M,\theta_e,\theta_u) L_f(\bar{t}\,|\,\cdot,\theta_e,\theta_u) dG(\theta_e,\theta_u) \ ,$$

where $G(\theta)$ denotes the heterogeneity distribution conditional on program eligibility. In (4), the contribution of the time remaining in the interrupted spells, $h(\cdot)$, is a complicated function of the probability of program eligibility, the entry rate into unemployment, and the hazard rate from unemployment (see Appendix B).[10] In light of these complications, we follow Heckman's and Singer's (1984a) suggestion and estimate an approximation to (4). In that approximation, we define a new hazard rate for leaving the interrupted unemployment spell conditional on program eligibility and time spent in the spell since the baseline, r,

$$(5) \qquad \lambda_r(r\,|\,\cdot,\theta_r) = \lambda_r(r\,|\,X_r,\theta_r;\beta_r) \ .$$

The survivor function corresponding to (5) is $S_r(r\,|\,\cdot\,,\theta_r)$. The contribution to the

[10]The interrupted spells' contribution can be simplified considerably by conditioning on the date, t^*, when these spells began. However, one must at the same time condition the heterogeneity distribution. Therefore, because the heterogeneity distribution for each control is $G^*(\theta\,|\,t^*)$ while for each trainee it is still $G(\theta)$, the random assignment in the experiment is contaminated.

likelihood for a control group member who leaves her interrupted spell after r periods and then experiences employment and unemployment spells, denoted by $\bar{t}$, is now given by

$$(6) \qquad L(r,\bar{t}\,|\,\cdot) = \int \lambda_r(r\,|\,\cdot,\theta_r) \cdot S_r(r\,|\,\cdot,\theta_r) \cdot L_f(\bar{t}\,|\,\cdot,\theta_e,\theta_u) \cdot dG(\theta_r,\theta_e,\theta_u),$$

where we have appropriately redefined the heterogeneity distribution function, $G(\theta)$. The contribution for a woman who never leaves her interrupted spell during the sampling frame is given by

$$(7) \qquad L(r\,|\,\cdot) = \int S_r(T\,|\,\cdot,\theta_r) dG_r(\theta_r).$$

In (7), $G_r(\theta)$ is the marginal distribution of θ_r, and T is the length of the sampling period after the baseline.

It is interesting to compare the controls' contribution to the likelihood function in our problem to the likelihood function for a typical sample selection problem, (e.g., Heckman (1979)). In (6) and (7), the contribution of controls' interrupted spells plays a role analogous to the probit equation in the standard sample selection problem. Likewise, the term for the fresh spells, $L_f(\cdot)$, plays the role of the regression equation. As in the sample selection problem, we may discard the interrupted spells and use the fresh spells to estimate the training effect only when the heterogeneity associated with the interrupted spells, θ_r, is independent of the heterogeneity associated with the fresh spells, θ_e and θ_u. Further, since some independent variables change with time, the interrupted spells' contribution to the likelihood is based on different values of those variables than the fresh spells' contribution to the likelihood. That difference should aid in the identification of the hazard function's parameters. Finally, we should emphasize that we do not correct for selection bias in this problem because we wish to obtain structural as opposed to reduced form parameter estimates (or unconditional versus conditional estimates). Instead, we correct for the selection process into fresh spells so that we can make the parameters of the treatments' and controls' hazard functions comparable.

4.3 The Contribution of the Training Spells

In contrast to the controls' interrupted spells, the treatments' unemployment spells end when they are randomly assigned into training (see Figure 2b). The treatments were eligible for up to S periods of training, but they could leave the program early if they dropped out or if they found a regular job. Therefore the treatments' contribution to the likelihood function must account for their time in training and whether or not they left training early for employment or unemployment. In order to model these possibilities, we define the transition intensity (see, e.g., Lancaster 1990) from training into unemployment after τ periods as

(8a) $$\lambda_{su}(\tau|\cdot,\theta_{su}) = \lambda_{su}(\tau|X_s,\theta_{su};\beta_{su}).$$

Similarly, we define the transition intensity from training into employment as

(8b) $$\lambda_{se}(\tau|\cdot,\theta_{se}) = \lambda_{se}(\tau|X_s,\theta_{se};\beta_{se}).$$

Given those definitions, the hazard function for exiting training is the sum of the two transition intensities

(9a) $$\lambda_{s}(\tau|\cdot,\theta_{su},\theta_{se}) = \lambda_{se}(\tau|\cdot,\theta_{se}) + \lambda_{su}(\tau|\cdot,\theta_{su}).$$

The survivor function corresponding to the hazard function in (9a) is given by

$$S_s(\tau | \cdot, \theta_{su}, \theta_{se}) = S_s(\tau | \cdot, \theta_{se}, \theta_{su}; \beta_{se}, \beta_{su}). \tag{9b}$$

To complete our specification of the training spells' contribution to the likelihood, we define a dummy variable δ_u that equals 1 when a woman drops out of training for unemployment and zero otherwise, and a dummy variable δ_e that equals 1 when a woman leaves training for employment and zero otherwise. The contribution of a training spell of t_s weeks to the likelihood function (conditional on the unobserved heterogeneity) is

$$L_s(t_s | \cdot, \theta_{su}, \theta_{se}) = \lambda_{se}(t_s | \cdot, \theta_{se})^{\delta_e} \cdot \lambda_{su}(t_s | \cdot, \theta_{su})^{\delta_u} \cdot S_s(t_s | \cdot, \theta_{su}, \theta_{se}). \tag{10}$$

To form the treatments' contribution to the likelihood, we combine the training period data with the fresh spell data and integrate out the heterogeneity to obtain

$$L(t_s, \bar{t}, \cdot) = \int_{\theta} L_s(t_s | \cdot, \theta_{su}, \theta_{se}) \cdot L_f(\bar{t} | \cdot, \theta_e, \theta_u) dG(\theta_{su}, \theta_{se}, \theta_e, \theta_u). \tag{11}$$

where we have again redefined the distribution function $G(\cdot)$.

As with the controls' contribution to the likelihood, we may draw an analogy from (11) to the sample selection literature. In this case, the contribution of the training spells, $L_s(\cdot)$, plays the role of the probit equation and the contribution of the fresh spells, $L_f(\cdot)$, plays the role of the regression equation. Accordingly, we may exclude the training spells from the analysis and still obtain consistent estimates of the training effect only when the heterogeneity terms associated with the training spells, (θ_{se}, θ_{su}), are independent of the heterogeneity terms from the fresh spells (θ_u, θ_e).

5. ESTIMATES OF THE TRAINING EFFECT

To estimate the effect of training on employment and unemployment duration, we maximize the likelihood formed by combining the controls' and treatments' contributions to the likelihood, given by equations (6), (7), and (11). This likelihood function is relatively complex and computationally demanding as it depends on five different parameter vectors (β_r, β_u, β_e, β_{su}, β_{se}) and five different heterogeneity terms, (θ_r, θ_u, θ_e, θ_{su}, θ_{se}).

The likelihood's complexity reflects the sample selection biases that arise when we exclude the controls' interrupted spells and the treatments' training spells. To motivate the use of our estimator and to indicate the potential importance of sample selection biases, we first present estimates of the training effects using only data from fresh spells. Next, we show how incorporating information on the controls' interrupted spells affects our estimates of the unemployment and employment hazards. Finally, we also include the treatments' training spells and estimate the complete likelihood function.

We first follow standard empirical practice and use only fresh employment and unemployment spells to estimate the training effect. By discarding the interrupted spells and training spells, we implicitly assume that the heterogeneity terms associated with those spells are independent of the heterogeneity terms in the employment and unemployment spells. In addition we also assume that the heterogeneity terms associated with the employment and unemployment spells are independent. Finally, we assume that the hazard function for employment and unemployment is given by

(13a)
$$\lambda_j(t_j|\cdot) = (1 + \exp(y_j))^{-1},$$

where

(13b)
$$y_j = \theta_j + \beta_j X(t_j + \tau_j) + \gamma_j D + \alpha_j \log(t_j) + \alpha_{2j} \log(t_j)^2,$$

where $j = e, u$.

The vector X in the foregoing hazard functions includes both personal characteristics and demand variables. Among the personal characteristics are age, years of schooling, whether or not the women dropped out of high school, the number of children less than 18, race, and marital status.[11] The demand variables are monthly nonagricultural employment and the number of persons receiving unemployment benefits. We measured both demand variables as log deviations from SMSA means. Finally, we used log duration and its square to capture the effects of duration dependence.[12]

The estimated training effects using only fresh spells suggest that training lowered the probabilities of leaving both employment and unemployment.[13] In Table 3, column (1) presents the coefficient (and standard error) of the training dummy variable and the duration terms from the employment hazard when θ_e is constant, while column (2)

[11]The means for treatments and controls characteristics are given in Table A of the Appendix. We also experimented with adding age-squared, and dummy variables for whether a woman was of Hispanic origin or currently married. None of these variables had a coefficient that was significantly different from zero, nor did the addition of the variables affect the results.

[12]The quadratic duration term was significant only in the employment hazard function. Consequently, we dropped it from the unemployment hazard. We also used time dummy variables instead of log duration and log duration squared to capture the effects of duration dependence (see Meyer 1989). This alternative specification had no effect on any of the estimated coefficients, including that for training status.

[13]The full set of parameter estimates are contained in Table B in the Appendix.

presents the same estimates when the heterogeneity term is assumed to come from a discrete distribution with two points of support.[14] The estimates indicate that training increased employment durations by approximately 11 months. In columns (3) and (4) of Table 4, we present the corresponding estimates for the unemployment hazard. The estimates indicate that training increased unemployment durations by approximately 40 months.[15]

In light of training's positive impact on employment rates, the finding that training impaired a woman's ability to find a job is surprising. However, a more plausible conclusion to draw from the analysis of fresh spells is that the sample selection problem is particularly serious when studying economically disadvantaged persons. The controls for unobserved heterogeneity in column (4) assume that the heterogeneity terms are independent of the explanatory variables, including training status. However, we argued above that using only fresh spells is likely to create a sample selection problem and to contaminate the experimental design. In this case, training status will be correlated with θ_u and θ_e and the standard approach is inappropriate.

An informal way of avoiding this selection problem is to include the controls' interrupted spells and to assume that the fresh and interrupted spells have the same hazard function.[16] But this procedure also involves the strong assumption of no duration

---

[14]See Heckman and Singer (1984b).

[15]The estimated differences in trainees' and controls' expected durations are calculated using the parameters from the estimated hazard functions, which ignore unobserved heterogeneity.

[16]We measure the duration in the spells as time spent unemployed since the baseline. Alternatively, if we use the actual duration of those spells, which includes time spent unemployed prior to the baseline, the estimated training effect is similar to that reported for fresh spells in column 3 of Table 3. By using actual duration, we adjust downward the hazard for controls' interrupted spells in order to take into account that

dependence in the unemployment hazard. If the assumption is inappropriate, including interrupted spells involves trading off the bias from misspecification — that is, treating time remaining in an interrupted spell the same as time in a fresh spell — against the sample selection bias resulting from excluding these interrupted spells.

As shown by column (5), when we use both interrupted and fresh unemployment spells, the estimated training effect in the unemployment hazard falls by one-half in absolute value and is no longer statistically significant at standard confidence levels. However, we also find strong evidence of duration dependence. Moreover, this duration dependence is not simply the result of neglected heterogeneity. As shown in column (6), when we assume that unobserved heterogeneity is drawn from a two-point distribution, the magnitude of the duration dependence coefficient declines but nonetheless remains substantial and statistically significant. The existence of duration dependence indicates that it is inappropriate to assume that the hazard function is the same for interrupted and fresh spells.

The findings in Table 3 indicate the potential importance of sample selection and motivate consideration of the more rigorous statistical framework developed in the previous section. To begin, we first allow the interrupted and fresh unemployment spells to have different hazard functions and require the heterogeneity terms in those functions to be correlated. We assume a one-factor loading structure where $\theta_r = r\theta_u$ and θ_u is drawn from a two-point distribution. Further, we assume that the heterogeneity terms

---

those women have been unemployed for some time at the baseline. If there were no unobserved heterogeneity, this alternative approach would be appropriate. But, in the presence of such heterogeneity, its distribution for interrupted spells should be conditioned for time spent unemployed prior to the baseline, again contaminating the experimental design.

associated with the interrupted and fresh unemployment spells are independent of the heterogeneity terms associated with the employment and training spells (θ_e, θ_{su}, θ_{se}). As shown in column (1) of Table 4, now training has essentially no effect on the transition rate out of a fresh unemployment spell.

We next consider the case where the heterogeneity terms from interrupted unemployment spells, fresh unemployment spells, and employment spells are correlated according to a one-factor structure where $\theta_r = r\theta_u$ and $\theta_e = e\theta_u$, and θ_u is drawn from a two-point distribution. Further, we assume that the heterogeneity terms associated with these spells (θ_r, θ_u, θ_e) are independent of the heterogeneity terms associated with the training spells (θ_{su}, θ_{se}). As shown by column (2), when we correct for selection bias in this fashion, training has no effect on unemployment durations. However, as shown by column (4), training continues to have a strong effect on employment durations.

Finally, we account for the potential selection bias arising from the treatments' exit from training. We assume that all the heterogeneity terms in our model are correlated and again follow a one-factor structure where $\theta_r = r\theta_u$, $\theta_e = e\theta_u$, $\theta_{su} = m\theta_u$, $\theta_{se} = n\theta_u$,[17] and θ_u is drawn from a two-point distribution.[18] Column (3) contains the coefficient for the training dummy from the unemployment hazard for the fresh spells, while column (5)

---

[17]We considered a generalization of this structure such that $\theta_{su} = \alpha_{su} + m\theta^*$, $\theta_{se} = \alpha_{se} + n\theta^*$, $\theta_e = \alpha_e + e\theta^*$, $\theta_r = \alpha_r + r\theta^*$ and $\theta_u = \alpha_u + \theta^*$ where θ^* is a *mean zero* random variable drawn from a two-point distribution. However, we tried this specification for a number of cases and starting values, and never had a nonnegligible effect on the likelihood value or the parameter estimates.

[18]The project was quite computationally demanding and thus we generally stayed with the assumption that θ_u was drawn from a two-point distribution. However, we did try to add a third point of support for our last model, which allows for full correlation. Even though we tried several starting values, we could only achieve a trivial increase in the likelihood. Given that we could not find a role for a third point of support in our most complicated model, we maintained the assumption that θ_u was drawn from a two-point distribution in our other specifications.

contains the estimate of the training coefficient from the employment hazard. Comparing columns (2) and (3), we see that the training dummy in the unemployment hazard has risen, but the estimated coefficient in column (3) still has an asymptotic normal statistic approximately equal to one. Comparing columns (4) and (5), we see that the training coefficient in the employment hazard has also risen and is quite significant.

Our estimates that account for sample selection indicate that training significantly increased the duration of employment spells, while it had no statistically significant effect on the duration of unemployment spells. We certainly found no evidence that training reduced unemployment durations. These findings make considerably more economic sense than the findings based only on fresh spells, which suggest that training substantially raised unemployment duration. The results also indicate that initial conditions problems are of more than theoretical interest in event history studies, and that policy conclusions based on these studies may be quite sensitive to how researchers deal with such problems.

VI. CONCLUSIONS

In this paper, we found that supported work raised employment rates because it helped women who found jobs remain employed longer than they would have otherwise. Our finding is in keeping with the program's objectives and is encouraging because longer employed spells may lead to greater human capital accumulation. Such a possibility suggests that the short-term program effects should persist and might even increase over time. A recent study by Couch (1991) supports this contention. Using social security quarterly earnings data, he reports that the NSW treatments had greater earnings than the controls more than 7 years after the supported work program ended. Thus, our study suggests that short sampling frames contain information that program evaluators might use to draw inferences about the long-term effects of training. Such a contention needs, of course, to be explored further in future research.

We conclude with a final point concerning the value of an experimental design when evaluating training's effect on employment and unemployment durations. The complexity of the estimator developed in this paper does not reflect a shortcoming of the experimental design. Indeed, in a nonexperimental setting this problem is much more complex. For example, Gritz (1988) uses the National Longitudinal Survey to evaluate the impact of public sector training on employment and unemployment durations. Because he does not have a random design, his study differs from ours in two fundamental ways.[19] First, since he does not condition the heterogeneity distribution for being eligible for training, his study addresses a more ambitious question than ours, namely, what effect training would have on a randomly chosen member of the labor force. Second, since he must allow for individuals entering training both before and during his sampling frame, he faces a more complex task in accounting for selection bias. Not surprisingly, Gritz finds that government-sponsored training substantially increases unemployment duration and decreases employment duration. He acknowledges that these findings may reflect the failure of his econometric model to account fully for selection bias.[20]

[19]He must also aggregate across different training programs.

[20]See Ridder (1986) for an evaluation of Dutch training programs with nonexperimental data. As he explicitly notes, Ridder is forced to make strong identifying assumptions since he lacks a control group and must instead rely on pre/post comparisons between unemployment durations.

In contrast to Gritz, we analyze the effect of training only among those who were eligible volunteers for the NSW program. Given the characteristics of individuals likely to participate in government-sponsored training programs, we do not see this as a serious limitation of our study. Moreover, for this group of eligible volunteers, random assignment assures that an individual's heterogeneity is independent of her training status. Because the experimental design eliminates the need to account for selection into training, we must simply model how the controls leave their interrupted unemployment spells and how the treatments leave their training spells. This task is clearly much more manageable than Gritz's and our results reflect this fact. Therefore, although the experimental design does not eliminate the need for a formal econometric model, it does give us sufficient leverage to obtain economically meaningful results.

APPENDIX A

DESCRIPTION OF NEW SUPPORTED WORK DATA

I. Source of Data and Documentation

The data used in this study were obtained from the Employment and Earnings File of the Supported Work Evaluation Study Public Use File. This file was prepared under contract number 33-36-75-01 to the Manpower Demonstration Research Corporation. The record layout and definitions of the variables in the public use file can be found in Technical Document No. 8 "Constructed Variables Derivation for the Supported Work Evaluation Study Public Use File: Employment and Earnings File," Mathematica Policy Research, Inc. and Social and Scientific Systems, Inc., December 1980. This paper uses data for the AFDC women who participated in the Demonstration.

II. Eligibility Requirements and Data Collection

To qualify applicants had to be currently unemployed, to have been unemployed for a total of at least 3 of the previous 6 months, to have received AFDC payments for 30 of the previous 36 months, and to have no preschool children. Eligible applicants who volunteered for Supported Work were randomly assigned into a treatment or a control group during 1976 and 1977. The experiment was run in seven sites: Atlanta, Georgia; Chicago, Illinois; Hartford, Connecticut; Newark, New Jersey; New York City, New York; Oakland, California; and in several locations in Wisconsin.

All participants, including the control group members, were interviewed when admitted into the program. Among the information collected in these interviews was the woman's age, years of schooling, whether she was a high school dropout, number of children under 18, marital status including whether she had ever been married, and race. In addition, retrospective data on a woman's employment status were obtained in semi-monthly intervals for the two years prior to the baseline. This information was used to calculate the respondent's number of semi-monthly periods of employment experience in the two years prior to the baseline. Another question determined the number of weeks since a woman's last regular job, which was used to construct a variable for whether a woman had held a regular job since she was 16 years old.

Both treatments and controls were interviewed at 9-month intervals following the baseline. These interviews collected information on each woman's employment status in semimonthly intervals during the previous nine months. These data were used to construct the length of spells of employment and unemployment during the 26 months following the baseline. Some women with a 27-month interval had employment data for only 26 months because their interview took place before the end of the month. The post-baseline employment histories in our study extend for 52 semi-monthly periods. The sample used in the study consists of only those women with a baseline and three 9-month interviews and who satisfied the two employment related eligibly criteria for the program. Unfortunately, less than 40 percent of the sample was interviewed after 27 months due to program costs. In addition, not every woman who participated in the program as either a treatment or a control appears to have satisfied the employment-related eligibility criteria. However, these factors do not affect the integrity of the experimental design since treatments and controls were affected equally. Nevertheless, the sample available for this study was greatly reduced. There were 275 women in our treatment group and 266 women in our control group. All of these women volunteered for the program during 1976. The means and standard errors of these women's demographic characteristics are presented in Table A.

The labor demand variables used in the paper were collected from various issues of Employment and Earnings published monthly by the U.S. Department of Labor. We used the deviation around the site mean of total payroll employment and number of persons receiving unemployment insurance to proxy for labor market conditions in each woman's city at a point in time.

III. Miscellaneous Issues

A. Deleting Ineligibles

There were thirty-four women — 19 trainees and 15 controls — whose employment histories prior to the baseline were inconsistent with two intended eligibility requirements of the program. Nearly all of these women were unemployed in less than 3 of the 6 months prior to the program; some also were employed at the baseline. When these women are put back into the sample there are 294 trainees and 283 controls. Ham and LaLonde (1990) presents the average durations and empirical survivor functions for this slightly larger sample. The program's effect on employment rates is unaffected by which sample we choose to use.

Excluding these women should not affect the integrity of the experimental design as long as "ineligible" women were no more likely to be assigned to the treatment group than to the control group. We focus on the program effects for this "eligible" sample in this paper partly because there are relatively few cases of ineligibles and because we would have too few data points to estimate a separate hazard for interrupted employment spells.

B. No-Shows

There were 14 treatment group members in our sample who volunteered and were randomly assigned into training but never showed up for supported work. We treat these no-shows as trainees throughout the analysis. To exclude these women from the analysis would contaminate the experimental design. Therefore, the training effect measures the impact on the employment opportunities of the treatment group members of the opportunity to participate in supported work.

C. Supported Work Participation

Trainees were guaranteed a subsidized supported work job for 9 to 18 months. Initially productivity and attendance standards for the participants were less than would be expected on a regular job, but these standards were raised by the program administrators over time. Some participants either left the program voluntarily or were asked to leave because of poor performance before their term expired. For the sample used in this paper, 75% of the participants had left by the 13th month and 95% were out of the program by the 17th month.

APPENDIX B

ANALYSIS OF THE PRE-BASELINE DATA AND OF THE CONTROLS' INTERRUPTED SPELLS

We face two problems in forming a likelihood that utilizes pre-baseline data. First, we must model the probability of being eligible for training, even if we are controlling for variation in unobserved variables only within the sample that volunteers for training. This necessitates obtaining a tractable expression for calculating this probability.

The second problem we face is that while we know the starting date of the unemployment spell in progress at the beginning of the pre-baseline data, we do not know the starting date of an employment spell in progress at the beginning of the pre-baseline data. We can write a likelihood for these data, but it is not clear how well some parameters will be identified. However we argue below that as a practical matter, it may be sufficient to know the starting date of the unemployment spell.

I. Specifying the Probability of Program Eligibility[21]

A woman is eligible for training if she was unemployed for *at least* six of the twelve semi-monthly periods preceding the baseline. The probability of spending at least six periods in unemployment in the interval (-12,0) depends on whether the individual is employed or unemployed at calendar time -12 and on the length of time she has been in employment or unemployment at this time.

---

[21]We owe a substantial debt to Geert Ridder for his help in deriving the expression for this probability.

We denote the probability of being eligible as $\Pr(M|\cdot)$ and define an indicator $I_u(-12) = 1$ if the individual is unemployed at time -12 and 0 otherwise. We note

(B.1) $$\Pr(M|\cdot) = \Pr(M, I_u(-12)=1|\cdot) + \Pr(M, I_u(-12) = 0|\cdot).$$

Define the probability of entering unemployment at time $-\tau_u$ as

$$k_u(-\tau_u | X_{ku}(\cdot), \theta_{ku}) ,$$

where θ_{ku} is an unobserved heterogeneity component.[22] Then

(B.2) $$\Pr(I_u(-12)=1|\cdot) = \int_0^\infty k_u(-(\bar{t}_u+12)|\cdot, \theta_{ku}) S_u(\bar{t}_u|\cdot, \theta_u) d\bar{t}_u$$

and

(B.3) $$\Pr(M, I_u(-12)=1|\cdot) = \int_0^\infty k_u(-(\bar{t}_u+12)|\cdot) S_u(\bar{t}_u|\cdot) \Pr(M|\bar{t}_u, \cdot) d\bar{t}_u .$$

Consider $\Pr(M|\bar{t}_u, \cdot)$, the probability of an individual meeting the eligibility criterion given that she has been unemployed for $\bar{t}_u$ periods at -12. Define N_u as the total number

[22] To obtain some intuition on the entry rate, consider the following example. Suppose in period 0, individuals are assigned to employment and unemployment by a fair coin toss. The probability of entering unemployment in period 3 is the employment hazard, λ_e, times the probability of being employed in period 2, $P(I_u(2)=0)$. Thus $k_u(3) = \lambda_e P(I_u(2)=0)$

$$= \lambda_e[(1-\lambda_e)P(I_u(1)=0) + \lambda_u P(I_u(1)=1)]$$

$$= .5\lambda_e[(1-\lambda_e)(\lambda_u+1-\lambda_e) + \lambda_u(\lambda_e+(1-\lambda_u))] .$$

The entry rate is a complicated nonlinear function of transition rates in previous periods but these cross-equation restrictions are not imposed in estimation.

of unemployment spells in [-12,0], *including* the spell interrupted at -12. Define N_e as the total number of employment spells in [-12,0]. In theory there could be a very large number of transitions in this interval, and N_e and N_u could be large integers. Even if we work in discrete time, there are a large number of employment histories over [-12,0] consistent with the eligibility criteria given $\bar{t}_u$. As a practical matter, one would expect to see very few transitions for this group. Thus while

$$\text{(B.4)} \quad \Pr(M \mid \bar{t}_u, \cdot) = \sum_{j=1}^{\infty} \Pr(M, N_u = j, N_e = j \mid \bar{t}_u, \cdot) + \sum_{j=0}^{\infty} \Pr(M, N_u = j+1, N_e = j \mid \bar{t}_u, \cdot),$$

one should obtain an excellent approximation by considering only the first few terms

$$\text{(B.5)} \quad \begin{aligned} \Pr(M \mid \bar{t}_u, \cdot) \approx\ & \Pr(M, N_u = 1, N_e = 0 \mid \cdot) + \Pr(M, N_u = 1, N_e = 1 \mid \cdot) \\ & + \Pr(M, N_u = 2, N_e = 1 \mid \cdot) + \Pr(M, N_u = 2, N_e = 2 \mid \cdot) \\ & + \Pr(M, N_u = 3, N_e = 2 \mid \cdot). \end{aligned}$$

One can calculate each of the probabilities in (B.5). For example,

$$\text{(B.6)} \quad \Pr(M, N_u = 1, N_e = 0 \mid \cdot, \bar{t}_u) = S_u(\bar{t}_u + 12 \mid \cdot, \bar{t}_u),$$

where $S_u(\bar{t}_u + 12 \mid \bar{t}_u)$ is the conditional probability of surviving for 12 additional periods given that the individual has survived up to $\bar{t}_u$. Further, one can impose the requirement that the individual be unemployed at the baseline by considering only the terms in (B.5) where $N_u > N_e$.

To calculate $\Pr(M, I_u(-12) = 1)$, first calculate (B.5), then substitute it into (B.3) for each value of $\bar{t}_u$ and integrate over this variable. We write

(B.8) $$\Pr(M, I_u(-12) = 1 \mid \cdot) = \Pr(M, I_u(-12)=1 \mid Z(\ ),\ \theta_{ku}, \theta_e, \theta_u)\ ,$$

where $Z(\ )$ is a vector of explanatory variables dating back to the time the individual left school. As a practical matter, it would be sensible to follow Nickell (1979) by first, working in discrete time and second, summing over a limited number of periods in the past.

A similar approach can be used to calculate the probability of being eligible for training for someone employed at -12,

(B.9) $$\Pr(M, I_u(-12) = 0 \mid \cdot) = \Pr(M, I_u(-12) = 0 \mid Z(\ ), \theta_{ke}, \theta_e, \theta_u)$$

where θ_{ke} is the unobserved heterogeneity component for the entry rate into employment. Using (B.1), the probability of being eligible for training is

(B.10) $$\Pr(M \mid \cdot) = \Pr(M \mid Z(\ ),\ \theta_{ke}, \theta_{ku}, \theta_e, \theta_u)\ .$$[23]

II. Contribution of a Control to a Likelihood Based on Pre- and Post-Baseline Data

There are 48 semi-monthly periods of (retrospective) data available prior to the baseline. Suppose that the individual is unemployed at period -48 and that this spell began at $\tau_u < -48$. (Note that τ_u is negative.) If we denote remaining duration in the spell by r_u, the contribution of this spell (conditional on θ_{ku} and θ_u) is

[23] John Mickelwright suggested that we treat the selection rule as unemployed at the baseline, leading to a simple stock sampling problem (as in Lancaster (1979) and Nickell (1979)). This would simplify matters considerably at the cost of misspecifying the selection rule.

(B.11) $$k_u(\tau_u \mid \cdot, \theta_{ku}) f_u(r_u - \tau_u - 48 \mid \cdot, \theta_u)\cdot$$

Subsequent employment and unemployment spells will contribute a standard multispell likelihood along the lines of (3), which we continue to write as $L_f(\cdot, \theta_e, \theta_u)$. The contribution of such a control based on pre- and post-baseline data, and conditioned on the unobserved heterogeneity, is

(B.12) $$L_{cu}(\cdot \mid \cdot, \theta_e, \theta_u, \theta_{uk}, \theta_{ek})$$

$$= [\Pr(M \mid \cdot, \theta_e, \theta_u, \theta_{ek}, \theta_{uk})]^{-1} k_u(\tau_u \mid \cdot, \theta_{ku}) f_u(r_u - \tau_u - 48 \mid \cdot, \theta_u) L_f(\cdot \mid \cdot, \theta_e, \theta_u).$$

The contribution of a control who is unemployed at time -48 and remains in this spell until the end of the sampling frame is (conditional on the unobserved heterogeneity)

(B.13) $$L_{cu}(\cdot \mid \cdot, \theta_e, \theta_u, \theta_{ku}, \theta_{ke}) = [\Pr(M \mid \cdot)]^{-1}\ k_u(\tau_u \mid \cdot) S_u(52 - \tau_u \mid \cdot),$$

recalling that there are 52 biweekly post-baseline periods.

The unconditional contribution of a control who is unemployed at -48 is[24]

(B.14) $$L^*_{cu} = \int L_{cu}(\cdot \mid \cdot, \theta_{ke}, \theta_{ku}, \theta_e, \theta_u) dG(\cdot),$$

where $dG(\cdot)$ is the density of $(\theta_{ke}, \theta_{ku}, \theta_e, \theta_u)$.

The contribution of a control who is in an employment spell at -48 is more complicated, since we do not know the start date of this employment spell. Suppose this

[24]Note that since we are not attempting to extrapolate beyond the population of volunteers, we do not integrate separately over $\Pr(M \mid \cdot)$ to allow for the selection rule to change the distribution of heterogeneity.

interrupted employment spell lasts r_e periods after time -48. The joint density of $I_u(-48)$ = 0 and r_e is

(B.15)
$$g_e(I_u(-48) = 0, r_e \mid Z(\ ), X_e(\ ), \theta_{ke}, \theta_e)$$
$$= \int_{-\infty}^{-48} k_e(\tau_e \mid \cdot, \theta_{ke}) f_e(r_e - \tau_e - 48 \mid \cdot, \theta_e) d\tau_e\ .$$

Again denoting contribution of further spells by $L_f(\cdot \mid \cdot, \theta_e, \theta_u)$, the contribution for this control is

(B.16)
$$L_{ce}^{*} = \int \left[\Pr(M \mid \cdot)^{-1} g_e(\cdot \mid \cdot, \theta_{ke}, \theta_e) L_f(\cdot \mid \cdot, \theta_e, \theta_u) d\bar{G}(\cdot) \right].$$

(The case where the individual remains in the employment spell over the sample period is uninteresting, since she will not be eligible for training.)

There are two troublesome features of the contribution to the likelihood for a control based on the pre- and post-sample data. First, calendar subscripts have been suppressed to avoid notational clutter, but the likelihood will depend on explanatory variables prior to time -48. In principle one can calculate the earlier values of these variables, particularly if one goes back only a small number of years to keep computational demands within reason.

Second, the entry rate for employment enters the likelihood only through integrals (or sums in discrete time), which may make identification of its parameters quite difficult. However, only 11% of the sample are in employment at -48, and only 4% of the sample are in employment at -12 (for calculating the probability of selection). Thus, as a practical matter, discarding the data on interrupted spells of employment at -48 may not lead to serious biases in the estimates. (Note that this affects trainees and controls to the same degree, as opposed to the case of the interrupted unemployment spells at the baseline.) Second, it may be appropriate to use the approximation

$$\Pr(M \mid \cdot) \approx \Pr(M, J_u(-12) = 1 \mid \cdot)\,. \tag{B.17}$$

III. Contribution of a Treatment to a Likelihood
Based on Pre-Baseline and Post-Baseline Data

The contribution for a treatment in the pre-baseline and post-baseline is derived in an analogous manner. Essentially one uses an expression such as (B.12) to describe the contribution (conditional on the heterogeneity) of her employment history up to the baseline. Multiply this expression times her contribution to the post-baseline data (i.e., equation (3) times equation (10)), and then integrate out the heterogeneity from this overall expression.

IV. The Contribution of Time Remaining in a
Spell Interrupted at the Baseline for a Control

First, we consider the case where there is no eligibility requirement for a control, and then we simply observe a control in an unemployment spell in progress at the baseline. In this case, the density of time spent in the spell after the baseline is given by

$$g(r \mid \cdot, \theta_{ru}, \theta_u) = \int_{\infty}^{0} k_u(-\tau_u \mid \cdot, \theta_{ru}) f_u(\tau_u + r \mid \cdot, \theta_u) d\tau_u . \tag{B.18}$$

(Recall that we cannot condition on the start date of the spell as this will contaminate the random assignment.) However, we must also take into account the fact that volunteers must satisfy the eligibility requirements, which will in turn affect the distribution of spells in progress at the baseline. Thus we have (conditional on the heterogeneity)

$$h(r \mid \cdot, M, \theta_{ke}, \theta_{ku}, \theta_u, \theta_e) = \frac{g(r, M \mid \cdot)}{Pr(M \mid \cdot)}, \tag{B.19}$$

where $g(r, M \mid \cdot)$ is the density of the joint event that an individual is eligible to participate in the program and then spends r periods after the baseline in the interrupted spell. Thus the controls' contribution to the likelihood based on multiplying (3) by (B.19) and integrating out the unobserved heterogeneity is essentially as complicated as that based on the pre-baseline data (B.14).

ACKNOWLEDGEMENT

We are grateful to John Abowd, David Card, Christopher Flinn, James Heckman, Joseph Hotz, Lawrence Katz, George Jakubson, Tony Lancaster, Angelo Melino, Robert Moffitt, Kevin M. Murphy, Thomas Mroz, Geert Ridder, Robert Topel and James Walker for very helpful discussions. Seminar participants at British Columbia, Chicago, Georgetown, Michigan, Northwestern, Pittsburgh, Stony Brook, Toronto, and Yale made useful suggestions. William Anderson, Lee Bailey, Susan Skeath, and especially Tan Wang provided excellent research assistance. The Social Science and Humanities Research Council of Canada, the Industrial Relations Section at Princeton University, and the Graduate School of Business at the University of Chicago generously supported this work. We emphasize that we alone are responsible for any errors.

REFERENCES

Ashenfelter, Orley (1978). "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics* 60, 47-57.

Ashenfelter, Orley, and Card, David (1985). "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs." *Review of Economics and Statistics* 67 (4), 648-660.

Barnow, Burt (1987). "The Impact of CETA Programs on Earnings: A Review of the Literature," *Journal of Human Resources* 22 (2), 157-193.

Bassi, Lauri (1983). "The Effect of CETA on the Post-Program Earnings of Participants," *Journal of Human Resources* 18 (Fall): 539-56.

Bell, Steven, Enns, J., and Orr, Larry (1987). "The Effects of Job Training and Employment on the Earnings and Public Benefits of AFDC Recipients: The AFDC Homemaker-Home Health Aid Demonstrations. Final Report. Washington, D.C.: Abt Associates.

Card, David, and Sullivan, Daniel (1988). "Measuring the Effect of CETA Participation on Movements in and out of Employment," *Econometrica* 56, 497-530.

Couch, Kenneth (1991). "Long-Term Earnings Effects of the National Supported Work Experiment: Evidence for the Youth and AFDC Target Groups." University of Wisconsin.

Flinn, Christopher J., and Heckman, James J. (1982). "Models for the Analysis of Labor Force Dynamics," in *Advances in Econometrics*, vol. 1, JAI Press, 35-95.

Friedlander, Daniel (1988). "Subgrouip Impact and Performance Indicators for Selected Welfare Employment Programs." New York: MDRC.

Gritz, Mark (1988). "The Impact of Training on the Frequency and Duration of Employment." Mimeographed. University of Washington,

Ham, John, and LaLonde, Robert J. (1990). "Using Social Experiments to Estimate the Effect of Training on Transition Rates," in J. Hartog, G. Ridder, and J. Theeuwes (eds.), *Panel Data and Labor Market Studies*. North-Holland: Elsevier Science Publications.

Heckman, James J. (1979). "Sample Selection Bias as a Specification Error," *Econometrica* 47, 153-61.

Heckman, James J., and Singer, Burton (1984a). "Econometric Duration Analysis," *Journal of Econometrics* 24, 63-132.

Heckman, James J., and Singer, Burton (1984b). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica* 47, 247-283.

Hollister, R., Kemper, P., and Maynard, R., eds. (1984). The National Supported Work Demonstration. (Madison: University of Wisconsin Press).

Lancaster, Tony (1979). "Econometric Methods for Duration of Unemployment," *Econometrica*, 47, 939-956.

______ (1990). *The Econometric Analysis of Transition Data*, Cambridge: Cambridge University Press.

Lancaster, Tony, and Nickell, Stephen (1980). "The Analysis of Re-Employment Probabilities for the Unemployed," *Journal of the Royal Statistical Society*, Series A, 143, 141-165.

Meyer, Bruce (1989), "Semi-parametric Estimation of Hazard Models," unpublished mimeo, Northwestern University.

Nickell, Stephen (1979). "Estimating the Probability of Leaving Unemployment," *Econometrica*, 47, 1249-66.

Ridder, Geert (1986). "An Event History Approach to the Evaluation of Training, Recruitment, and Employment Programs," *Journal of Applied Econometrics*, 1.

Ridder, Geert, and Verbakel, Wim (1984). "On the Estimation of the Proportional Hazards Model in the Presence of Unobserved Heterogeneity," Report 17/84, Faculty of Actuarial Science and Econometrics, University of Amsterdam.

Wolthagen, C., and Goldman, Barbara (1983). "Job Search Strategies: Lessons from the Louisville WIN Laboratory." New York: MDRC.

TABLE 1

EMPIRICAL SURVIVOR FUNCTIONS

[Proportion Remaining Employed or Unemployed]

	Employment		Unemployment		
Months	Treatments	Controls	Treatments	Controls: All Spells	Controls: Fresh Spells
1/2	0.968	0.929	0.955	0.949	0.929
	(.013)	(.018)	(.013)	(.011)	(.023)
1	0.929	0.848	0.910	0.914	0.895
	(.019)	(.026)	(.018)	(.015)	(.028)
2	0.839	0.761	0.864	0.843	0.791
	(.027)	(.030)	(.021)	(.019)	(.039)
3	0.787	0.687	0.817	0.807	0.756
	(.031)	(.033)	(.024)	(.021)	(.039)
4	0.733	0.648	0.778	0.781	0.728
	(.013)	(.018)	(.013)	(.011)	(.023)
5	0.670	0.603	0.746	0.756	0.672
	(.013)	(.018)	(.013)	(.011)	(.023)
6	0.650	0.573	0.730	0.725	0.613
	(.013)	(.018)	(.013)	(.011)	(.023)

Notes.--The calculations in Column 4 include spells in progress at the baseline. (In the spells in progress, duration is measured from the baseline.) Those in Column 5 use only unemployment spells that begin after the baseline. The standard error calculations account for "right censoring" of the data.

TABLE 2

INDIVIDUAL AND SPELL CHARACTERISTICS

Variable	Employment: Treatments	Employment: Controls	Unemployment: Treatments	Unemployment: Controls: All Spells	Unemployment: Controls: Fresh Spells
Panel A: Spell Characteristics					
All Spells:					
Mean Duration	14.98 (0.79)	15.67 (1.04)	21.18 (0.86)	28.63 (1.05)	15.75 (1.26)
Number of Spells	185	198	269	374	126
Completed Spells:					
Mean Duration	8.61 (0.79)	9.75 (1.04)	11.09 (0.86)	15.57 (1.05)	9.87 (1.26)
Number of Spells	81	121	107	185	61
Panel B: Individual Characteristics					
Age	33.77 (.60)	34.73 (.63)	33.21 (.51)	34.80 (.44)	34.36 (.72)
Schooling	10.42 (.14)	10.55 (.17)	10.18 (.13)	10.11 (.13)	10.50 (.20)
H.S. Dropout	.62 (.04)	.60 (.04)	.71 (.03)	.71 (.03)	.63 (.05)
Kids under 18	2.29 (.10)	2.41 (.12)	2.26 (.09)	2.30 (.08)	2.40 (.15)
Never Married	.38 (.04)	.30 (.04)	.39 (.03)	.33 (.03)	.34 (.05)
Proportion Black	.83 (.03)	.78 (.04)	.86 (.02)	.83 (.02)	.82 (.04)
Prior Experience	2.46 (.63)	4.09 (.65)	2.82 (.40)	2.83 (.40)	5.04 (.87)
Proportion Never Employed	.17 (.03)	.13 (.03)	.15 (.02)	.18 (.02)	.12 (.03)
Number of Women	149	138	222	266	92

Notes.--Mean durations are the mean number of semi-monthly periods. All employment spells and trainees' unemployment spells begin after the baseline. The controls' spells in column 4 include both unemployment spells that are in progress at the baseline and that begin after the baseline. Duration of those spells in progress at the baseline is measured as semi-monthly periods from the baseline. The statistics in column 5 include only spells that begin after the baseline. The numbers in parentheses are the standard errors. Prior experience is number of semi-monthly periods worked in the two years preceding the baseline.

TABLE 3

ESTIMATED TRAINING EFFECTS USING FRESH SPELLS

	Employment		Unemployment			
			Fresh Spells Only		Fresh and Interrupted Spells Combined	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Training Status	-.394	-.425	-.382	-.374	-.191	-.105
	(.155)	(.180)	(.166)	(.208)	(.129)	(.156)
Log Duration	.212	.155	-.453	-.299	-.503	-.353
	(.246)	(.258)	(.075)	(.119)	(.055)	(.082)
Log Duration Squared	-.168	-.113	---	---	---	---
	(.069)	(.085)				
Controls for Unobserved Heterogeneity	No	Yes	No	Yes	No	Yes
Log Likelihood	-852.7	-851.6	-768.1	-765.5	-1416.4	-1411.9

Notes: All models include controls for age, years of schooling, a woman's high-school dropout status, number of children under 18, marital status, race, and SMSA establishment employment and unemployment insurance recipients. The standard errors are in parentheses.

TABLE 4

ESTIMATES OF THE TRAINING EFFECT BASED ON ALTERNATIVE HETEROGENEITY ASSUMPTIONS

	Fresh Unemployment Hazard			Employment Hazard	
	(1)[a]	(2)[b]	(3)[c]	(4)[b]	(5)[c]
Training Status	-0.0800 (0.2588)	-0.0103 (0.1948)	-0.2090 (0.2000)	-0.2809 (0.1828)	-0.4470 (0.1680)
Log Likelihood[d]	-1404.6	-2251.7	-3000.3	-2251.7	-3000.3
Log Likelihood[d] No Heterogeneity	-1407.9	-2260.7	-3005.7	-2260.7	-3005.7

[a]Assuming only θ_r and θ_u are correlated.

[b]Assuming θ_r, θ_u, and θ_e are correlated.

[c]Assuming all heterogeneity terms are correlated.

[d]In column (1) the log likelihood refers to the contribution of the fresh and interrupted unemployment spells. In columns (2) and (4) it refers to the contribution from employment spells as well as that from the interrupted and fresh unemployment spells. In columns (3) and (5) the log likelihood refers to the contribution of all spells (i.e., including the training data).

TABLE A

CHARACTERISTICS OF TREATMENTS AND CONTROLS

	Treatments	Controls
Age	33.67 (.47)	34.98 (.45)
Schooling	10.23 (.12)	10.18 (.13)
H.S. Dropout	.70 (.03)	.71 (.03)
Number of Kids	2.25 (.08)	2.31 (.08)
Never Married	.38 (.02)	.33 (.03)
Black	.85 (.02)	.83 (.02)
Prior Experience[1]	2.59 (.34)	2.91 (.41)
Never Employed[2]	.16 (.02)	.18 (.02)
Number of Women	275	266

[1]Prior experience is the number of semi-monthly periods of employment in the two years prior to the baseline.

[2]Never employed is a dummy variable indicating that the woman has not had a regular job since she was 16 years old.

TABLE B

ESTIMATES OF EMPLOYMENT AND UNEMPLOYMENT HAZARD FUNCTIONS

A: Full Set of Estimates for Table 3

	Employment Spells		Fresh Unemployment Spells		All Unemployment Spells	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Training status	-.394	-.425	-.382	-.374	-.191	-.105
	(.155)	(.180)	(.166)	(.208)	(.129)	(.156)
Age	-.013	-.016	-.009	.009	-.016	-.018
	(.011)	(.013)	(.012)	(.015)	(.009)	(.011)
Schooling	.049	.056	.125	.149	.047	.169
	(.051)	(.062)	(.063)	(.078)	(.045)	(.054)
H.S. Dropout	.398	.472	-.306	-.427	-.335	-.437
	(.199)	(.238)	(.219)	(.277)	(.163)	(.200)
Kids under 18	.010	.009	-.037	-.024	.030	.052
	(.055)	(.066)	(.060)	(.073)	(.045)	(.053)
Never Married	.075	.028	-.148	-.182	-.243	-.284
	(.167)	(.205)	(.182)	(.238)	(.134)	(.167)
Black	.060	.065	-.391	-.458	-.386	-.469
	(.191)	(.225)	(.210)	(.275)	(.153)	(.191)
Area Employment	1.78	.110	4.60	3.36	.531	1.15
	(3.78)	(4.18)	(4.24)	(4.83)	(3.10)	(3.38)
Area Unemployment	.651	.669	-.117	-.135	-1.26	-1.19
	(.431)	(.447)	(.511)	(.531)	(.368)	(.385)
Log Duration	.212	.155	-.453	-.299	-.503	-.353
	(.246)	(.258)	(.075)	(.119)	(.055)	(.082)
Log Duration^2	-.168	-.113	---	---		
	(.069)	(.845)				
θ_1	-3.15	-2.78	-2.91	-2.21	-3.16	-2.34
	(.817)	(1.03)	(.946)	(1.36)	(.694)	(.885)
θ_2	---	-4.12	---	-3.81	---	-3.94
		(1.19)		(1.36)		(.862)
μ	---	.181	---	1.24	---	1.68
		(1.20)		(1.50)		(.836)
-Log L	852.7	851.6	768.1	765.5	1416.4	1411.9

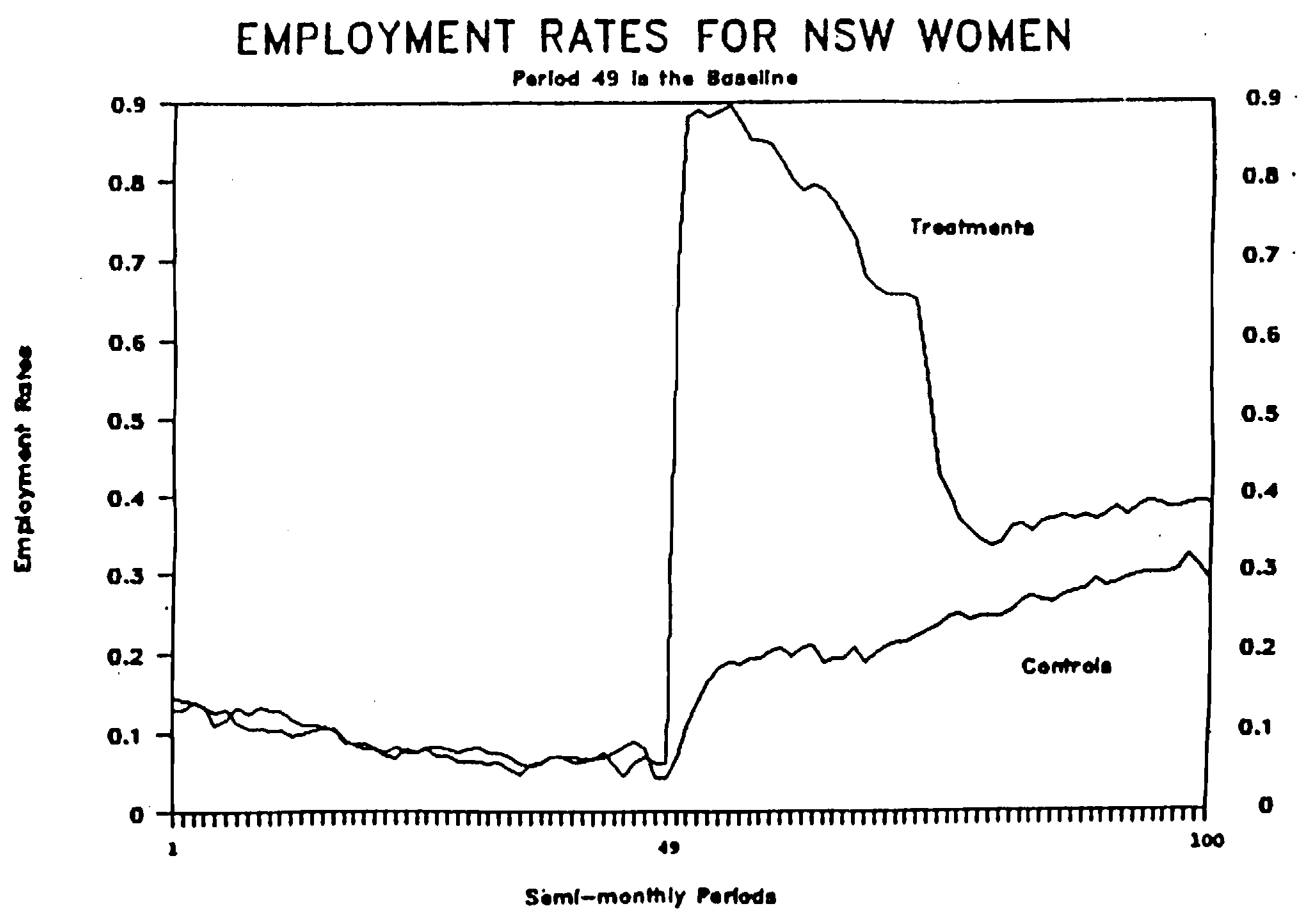
TABLE B (cont'd)

B: Full Set of Estimates for Table 4

Variable	Excluding Training and Employment Spells: Unemployment: Interrupted (1)	Fresh (2)	Excluding Training Spells: Unemployment: Interrupted (3)	Fresh (4)	Employment (5)	Including Training Spells: Unemployment: Interrupted (6)	Fresh (7)	Employment (8)
Training status	---	-0.0800	---	-0.0103	-0.2809	---	-.209	-.447
		(0.2588)		(0.195)	(.183)		(.200)	(.168)
Age	-0.0300	-0.0092	-0.0330	-0.0118	-0.0098	-.027	-.007	-.016
	(0.0143)	(0.0144)	(0.0147)	(0.0127)	(0.0122)	(.014)	(.015)	(.012)
Years of schooling	0.1553	0.1710	0.1307	0.1621	0.0617	.168	.146	.020
	(0.0764)	(0.0770)	(0.0601)	(0.0663)	(0.0585)	(.073)	(.084)	(.060)
Kids under 18	0.1736	-0.0296	0.1547	-0.0467	0.0049	.148	-.039	.019
	(0.0795)	(0.0737)	(0.0791)	(0.0655)	(0.0620)	(.071)	(.067)	(.060)
Black	-0.5170	-0.4120	-0.4412	-0.3485	0.0769	-.478	-.386	.001
	(0.2638)	(0.2946)	(0.2673)	(0.2309)	(0.2092)	(.240)	(.237)	(.205)
Never married	-0.3240	-0.2193	-0.3498	-0.2134	0.0695	-.329	-.109	.075
	(0.2272)	(0.2316)	(0.2307)	(0.2003)	(0.1830)	(.211)	(.222)	(.182)
H.S. dropout	-0.5503	-0.4520	-0.5744	-0.2419	0.4664	-.472	-.326	.292
	(0.2706)	(0.2822)	(0.2578)	(0.2343)	(0.2230)	(.253)	(.225)	(.243)
Local employment	-4.2014	3.1148	-2.8183	2.5632	-0.6375	-4.168	4.697	2.205
	(5.4625)	(5.1231)	(5.3823)	(4.4081)	(3.9943)	(5.308)	(4.393)	(3.844)
Local unemployment	-2.5277	-0.1334	-2.5591	-0.0893	0.6708	-2.450	-.150	.626
	(0.5771)	(0.5355)	(0.5738)	(0.5261)	(0.4440)	(.571)	(.516)	(.438)
Log duration	-0.4746	-0.2779	-0.5002	-0.3459	0.1044	-.511	-.396	.184
	(0.1277)	(0.1306)	(0.1114)	(0.0845)	(0.2506)	(.114)	(.093)	(.246)
Log duration squared	---	---	---	---	-0.0998	---	---	-.150
					(0.0751)			(.069)
θ_1	-0.0282		-0.0605			-.037		
	(0.0135)		(0.0181)			(.013)		
θ_2	-0.0451		-0.0335			-.026		
	(0.0163)		(0.0096)			(.092)		
μ	-1.0028		-0.5257			2.033		
	(1.2078)		(0.3694)			(1.771)		
LOADING	0.7433	1.0000	0.6153	1.000	0.9911	.889	1.000	.778
	(0.3795)		(0.2772)		(0.3749)	(.449)		(.520)
Log L	-1404.6		-2251.7			3000.3		
Log L, No heterogeneity	-1407.9		-2260.5			3005.7		

NOTES.--Local employment and unemployment are defined as the log deviations from site means. The term μ is defined as $P_1 = e^{\mu}/(1+e^{\mu})$. The "loading" factors are defined as $j = r, u, e$ where $\theta_j = j\theta_1$ and $j = u$ is normalized to equal 1.

Figure 1



EMPLOYMENT RATES FOR NSW WOMEN
Period 49 is the Baseline
Employment Rates
Treatments
Controls
Semi-monthly Periods

FIGURE 2

TREATMENTS AND CONTROLS LABOR MARKET HISTORIES

Figure 2A: Controls

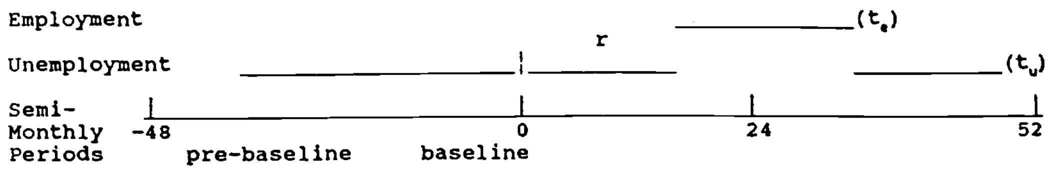
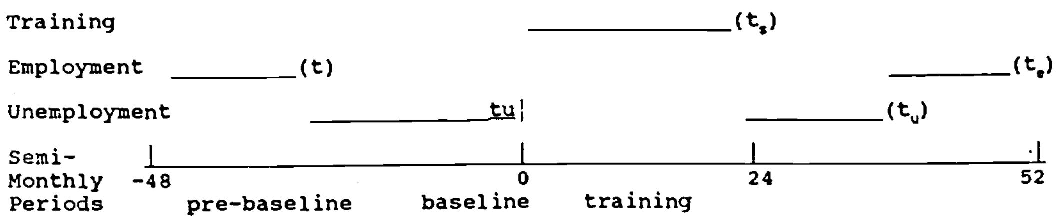


Figure 2B: Treatments



Notes: Time is measured in semi-monthly periods. Eligible women volunteer and are randomly assigned into the treatment or control group at the baseline -- time 0. The treatment group members leave their spell of unemployment and receive training for approximately 1 year (24 periods) while the controls continue in their spell of unemployment.