The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training
Author(s): John C. Ham and Robert J. Lalonde
Reviewed work(s):
Source: Econometrica, Vol. 64, No. 1 (Jan., 1996), pp. 175-205
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/2171928
Accessed: 14/03/2013 13:21

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.
We investigate the separate effects of a training program on the duration of participants' subsequent employment and unemployment spells. This program randomly assigned volunteers to treatment and control groups. However, the treatments and controls experiencing subsequent employment and unemployment spells are not generally random (or comparable) subsets of the initial groups because the sorting process into subsequent spells is very different for the two groups. Standard practice in duration models ignores this sorting process, leading to a sample selection problem and misleading estimates of the training effects. We propose an estimator that addresses this problem and find that the program studied, the National Supported Work Demonstration, raised trainees' employment rates solely by lengthening their employment durations.

KEYWORDS: Duration models, job training, sample selection.

1. INTRODUCTION

GOVERNMENT SPONSORED EMPLOYMENT and training programs frequently improve the labor market prospects of economically disadvantaged women. Often, 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 (Gueron and Pauly (1991)). Less is known, however, about how these programs generate these employment gains. Training may raise employment rates because it helps unemployed former participants find jobs faster by improving their job search skills. Alternatively, training may be effective because it helps employed participants retain regular jobs by improving their work habits. Because the

1 We are grateful to John Abowd, David Card, Christopher Flinn, Joseph Hotz, Lawrence Katz, George Jakubson, Kris Jacobs, Tony Lancaster, Angelo Melino, Bruce Meyer, Robert Moffitt, Kevin M. Murphy, Thomas Mroz, Robert Porter, Joseph Tracy, Robert Topel, James Walker, and two anonymous referees for helpful discussions and comments. We owe an especially large debt to James Heckman, Bo Honore, and Geert Ridder. Seminar participants at British Columbia, Chicago, Georgetown, Michigan, Northwestern, Pittsburgh, Stony Brook, Toronto, Virginia, and Yale made many useful suggestions. William Anderson, Lee Bailey, Susan Skeath, and especially Kris Jacobs and 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, and NSF (SES-9213310) generously supported this work. Part of this research was carried out while Ham was a visitor in the Economics Department at Northwestern and he would like to thank the department for its support and hospitality. We emphasize that we alone are responsible for any errors.

2 Both experimental and nonexperimental evaluations of government sponsored training administered during the past three decades consistently find that these programs raise the earnings and employment rates of adult women. By contrast, these evaluations usually find that these programs do not benefit adult men and youths. For surveys of these studies see Barnow (1987), Gueron and Pauly (1991), and LaLonde (1995).
services provided by government-sponsored training programs vary widely, we would not be surprised to find that these programs have varying impacts on employment and unemployment durations. To better understand how training raises employment rates, we develop an econometric framework for estimating the separate effects of training on the durations of participants’ subsequent employment and unemployment spells.

There are several reasons why such estimates should interest policy makers and analysts. First, by distinguishing between training’s impacts on employment and unemployment durations, our estimates enable us to look into the “black box” and learn how training works. This distinction is helpful to analysts interested in understanding training’s effects within the context of theoretical models of job search behavior, firm hiring decisions, and employee turnover.

A second and more immediate benefit of our estimates is that they may aid in the design of new training programs. Policy makers generally would prefer to combine a service that helps participants find jobs with one that helps them hold on to their jobs, as opposed to combining two services that each help trainees leave unemployment. Alternatively, policy makers may prefer to fund a program that lengthens employment durations as opposed to one that shortens unemployment durations, because the former program is likely to lead to more stable job histories and greater human capital accumulation. Unfortunately, it is not possible to make such policy decisions if we only know the effect of training programs on employment rates.

A final benefit of our approach is that we can use our estimates to predict the effect of a program beyond the sampling frame. We can do this by simulating our econometric model or by calculating the steady state employment rate (the ratio of the expected duration of an employment spell to the sum of the expected durations of an employment spell and an unemployment spell). This application should be useful to program evaluators because trainees usually are followed for only one or two years after leaving the program and there is reason to believe that the long-run effects of many programs differ from their short-run effects.

In this paper we apply our econometric framework to data on disadvantaged women from a social experiment. The advantage of data from a randomized experiment is that among the population of eligible program volunteers, a woman’s training status is uncorrelated with her unobserved heterogeneity. Therefore, simple comparisons between trainees’ and controls’ employment rates yield unbiased estimates of training’s effect on the probability of employment. However, a similar comparison between the durations of trainees’ and controls’ employment and unemployment spells, or their hazard rates out of those spells, yields potentially biased and economically misleading estimates of the effect of training. Although program administrators used random assignment to create the treatment and the control groups, there is no reason to believe that the treatments and controls experiencing subsequent employment and unemployment spells are random subsets of the experimental sample. In fact, we present strong evidence below that the samples of individuals in new spells are not random subsets of the experimental sample. Consequently, even
when using experimental data, evaluations of training's effect on employment and unemployment durations require a formal statistical model.

Our empirical findings may be summarized as follows. First, we find that using only new employment and unemployment spells creates a sample selection problem that contaminates the experimental design and thereby yields misleading estimates of the training effects. Second, our econometric framework successfully addresses this sample selection problem. This finding is important because there is no practical modification to the experimental design that would eliminate this problem in our data. Third, although this problem requires a formal statistical model, random assignment provides crucial identifying information in a relatively complex econometric model. Finally, the social experiment studied in the paper—the National Supported Work (NSW) Demonstration—raised trainees' employment rates because it helped those who found jobs to keep them. This program had no effect on the rate at which individuals left unemployment.

The remainder of the paper proceeds as follows. Section 2 discusses the problems that occur when experimental data are used to make inferences about the effect of training on employment and unemployment durations. Section 3 focuses explicitly on the NSW program to illustrate the issues discussed in Section 2. Section 4 constructs an econometric model that formally addresses the problems raised in the previous sections. Section 5 reports our empirical findings. Section 6 concludes the paper.

2. DURATION ANALYSIS WITH EXPERIMENTAL DATA

An appealing feature of social experiments is that they yield easily derived estimates of a program's impact on a variety of policy-related outcomes. In these experiments, program administrators randomly assign eligible applicants either into a treatment group—whose members are offered program services—or into a control group—whose members are denied access to these services. A woman's experimental status is, by construction, independent of her other characteristics, and thus the difference between the treatment and control groups' mean employment rates or earnings provide an unbiased estimate of the program's impact. Additional controls for differences among women's characteristics do not affect (asymptotically) the estimated impact although they can affect its standard error.

Alternatively, when a program does not incorporate an experimental design, the evaluation of its impact becomes much more complex. In such a setting, researchers first must construct a comparison group of persons who did not participate in the program, and then specify a statistical model that simultaneously accounts for the selection process into training and for the process that generates the outcomes of interest (Ashenfelter and Card (1985), Heckman and Robb (1985), Card and Sullivan (1988)). In practice, the estimated outcomes of such programs have been sensitive to how researchers constructed their comparison groups and how they specified their econometric models. Indeed, the sensitivity of nonexperimental estimates has generated a debate about their

However, when we turn from the question of whether training is effective to the question of how or why it works, even with experimental data we often must rely on nonexperimental methods like those described above. Consider, for example, a typical employment and training program for disadvantaged individuals in which volunteers are unemployed when they are randomly assigned to training at the baseline. At that time, the treatments leave unemployment and enter training for a period of $t^*$ weeks. Afterwards, program evaluators follow their progress in the labor market for an additional $T - t^*$ weeks. By contrast, the controls remain unemployed at the baseline and then are followed for a total of $T$ weeks. Suppose that the subsequent experimental evaluation demonstrates that training significantly raised participants’ earnings in week $T$. Assume for expositional purposes that training did not affect weekly hours of work (conditional on employment status) in a given period. Instead, it achieved these earnings gains (i) by raising employment rates and/or (ii) by raising hourly wage rates, and we wish to estimate these two effects.

Measuring the effect of training on employment rates is straightforward. Program evaluators obtain an unbiased estimate of the average effect during week $T$ simply from the difference between the treatment and control groups’ employment rates. Likewise, a seemingly intuitive way to measure the effect of training on wages is to compare the two experimental groups’ mean wages during week $T$. However, if training affected the treatments’ employment rates, this simple estimator of the wage effect is biased. This bias arises because we observe wages only for the employed, and because an individual’s employment status depends on his or her experimental status as well as other observed and unobserved characteristics.

To see how this bias arises, suppose further that individuals differ in only two respects: their training status and whether or not they are high school dropouts. In addition, assume that an individual’s dropout status is not observed by the econometrician and that dropouts have lower employment rates and lower wages than high school graduates. Suppose that when training raises employment rates, a larger fraction of trainees who are dropouts find jobs than is the case for controls who are dropouts. Because the employed trainees are (on average) less educated, the difference between their mean wages and those of the employed controls underestimates the effect of training on wages. This bias arises because the outcome of interest, hourly wages, is missing for some individuals and whether it is missing depends on an individual’s experimental status and unobservables (dropout status). As a result, training status is (negatively) correlated with the unobservables in the sample of employed persons, and this correlation prevents us from obtaining an unbiased estimate of the average treatment effect by simply comparing the mean hourly wages of the employed treatments and controls.

It is worth observing in our hourly wage example that (i) there is no correction to the experiment that will solve this missing data problem and (ii) this problem
SAMPLE SELECTION AND INITIAL CONDITIONS

does not imply that random assignment is of no use in examining the effect of training on hourly wages. To estimate this effect with experimental data, we must model the selection process into employment using a nonexperimental approach such as that suggested in Heckman (1979). By contrast, to estimate the wage effect with nonexperimental data, we must model not only (i) the selection process into employment but also (ii) the (more difficult) selection process into training. Consequently, the availability of experimental data significantly simplifies the task of estimating the effect of training on hourly wages. 3

A similar but more complicated problem arises in our example when we examine how training increased employment rates. In this case, we wish to observe the effect of training on the length of new employment and unemployment spells. Consider first the effect of training on employment spells. A natural way to proceed is to compare the treatments' and controls' experiences in these spells. However, there is no reason to believe that during the sampling frame we will observe comparable fractions of treatments and controls who are dropouts in employment spells. On the one hand, training may make it relatively easier for treatments to leave unemployment in a given period, especially for treatments who are high school dropouts. This would tend to increase the fraction of treatments in new employment spells who are dropouts relative to members of the control group that experience these spells. On the other hand, the trainees have only \( T - t^* \) periods to find a job after they leave training, whereas the controls have all \( T^* \) periods to find employment, and this difference may lead to a higher fraction of controls being dropouts in employment spells. Unless these effects cancel out, the fraction of employed treatments and controls who are dropouts will differ, and training status will be correlated with unobservables (dropout status). As a result, random assignment will be contaminated in the sample of women experiencing employment spells, just as it was contaminated in the comparison of hourly wage rates.

It is interesting to ask whether we can say training “causes” wages in the sense of Holland (1986), Rubin (1986), and the papers cited therein. It is clear that interpreting training as causing employment status appears consistent with this literature, because offering individuals training can be (i) manipulated (Rubin, p. 962) and (ii) offered as a treatment in an experiment (Holland, p. 954). Our example of the effect of training on wages differs from Holland’s discussion in that he assumes that the outcome of interest is observed for each individual, while in our example, wages are missing nonrandomly. Therefore, even though offering training is a manipulable treatment in an experiment, it is not in general possible to construct an experiment to simply measure the effect of training on wages, and it is not clear whether one should describe training as causing wages in Holland’s sense. We leave this as an open issue and continue to use the terminology of the “effect of training on wages” and “the effect of training on unemployment duration” in what follows. We note that Holland’s and Rubin’s use of the word “cause” is narrower than that used in much of economics. For example, it probably would not be consistent with Holland’s paper to speak of the effect of time spent studying on grades or test scores. The problem is that time spent studying can not be directly manipulated, even though it may be indirectly affected by randomly assigning individuals to a treatment group that (i) received computers or (ii) was assigned to additional mandatory study periods. (See Holland’s discussion on pp. 954–955.) We are grateful to an anonymous referee for drawing our attention to the Holland paper.
Direct comparisons between the treatments' and controls' unemployment spells may involve even more serious biases. The treatments' unemployment spells are truncated at the baseline when they enter training, and they begin fresh or new unemployment spells when they finish training at $t^*$. By contrast, after random assignment the controls remain in their unemployment spells and continue to look for a regular job. For a control to experience a fresh unemployment spell, she must leave the unemployment spell in progress at the baseline to begin a new job, and then subsequently leave employment for unemployment within the sampling frame. However, in a sample of disadvantaged persons, many controls will not leave the unemployment spell in progress at the baseline during the sampling frame. (In what follows we will refer to these spells as interrupted.) Further, among those controls who leave this unemployment spell, some will not leave their subsequent fresh employment spell. As a result, the number of controls with fresh unemployment spells will be much smaller than the number of treatments with such spells. More importantly, if dropout status does not have a large effect on employment duration, it is likely that the controls who experience fresh unemployment spells will be predominantly those with high school degrees, since these controls are more likely to exit their interrupted unemployment spell.

If we could treat remaining time (after the baseline) in an interrupted unemployment spell as equivalent to time in a new unemployment spell (that began after the baseline), we could use the interrupted spells to avoid the missing data problem for the controls' spells. However, as we discuss in some detail below, time remaining in interrupted spells is not comparable to time spent in new spells if there is (i) duration dependence and (ii) unobserved heterogeneity. We find that both of these conditions hold in our data and believe that they are likely to hold when evaluating other employment and training programs targeted to economically disadvantaged persons. To address the contamination of random assignment in the samples of those experiencing fresh employment and unemployment spells, we must model the treatments' and controls' entry into such spells.

In the discussion above, we have ignored an additional problem that arises even if the fractions of high school dropouts and graduates were identical among the treatments and controls experiencing fresh unemployment (employment) spells. In practice, training takes time, sometimes as much as one year. As a result, some of the controls' fresh spells take place when the treatments are in training. Therefore, the treatments' and controls' spells will correspond to different local demand conditions, violating the premise underlying experimental evaluations that a women's training status is independent of her other characteristics. We could avoid this problem by comparing the probability that treatments and controls leave a fresh unemployment (or employment) spell during the same calendar week. However, we then would face the difficulty that because the treatments and controls started these spells at different calendar times, duration dependence and unobserved heterogeneity ensures that these exit probabilities will differ in general even in the absence of a treatment effect. We do not focus on this complication because (at least in principle) it may be
SAMPLE SELECTION AND INITIAL CONDITIONS

addressed using standard econometric models for transition data (Lancaster (1990)), whereas these models cannot deal with unobserved differences between the samples of treatments and controls who enter fresh employment and unemployment spells. Below we show that the standard approach can be very sensitive to the latter problem.

We should emphasize that even though we must use nonexperimental methods to address these questions, our task is made much easier when program administrators randomly assign eligible applicants into either a treatment or a control group. In any nonexperimental analysis of the effect of training on the duration of employment and unemployment spells, we would have to simultaneously model (i) the nonrandom selection process into training and (ii) the nonrandom selection into fresh employment and unemployment spells. Because we have experimental data we need only address the second of these problems. Given that we believe that it is very difficult to model the selection into training, we view the availability of experimental data as extremely beneficial.

3. THE EFFECT OF TRAINING ON EMPLOYMENT HISTORIES

To illustrate the potential importance of the points raised in the previous section, we examine training's impact on participants in the National Supported Work (NSW) demonstration. This program provided work experience to a random sample of eligible AFDC women who volunteered for training. The remaining volunteers did not receive training and thus formed the control group. Women in the treatment group usually were guaranteed 12 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 jobs ended, the trainees were expected to enter the labor market and find regular jobs.

Despite similar preprogram employment rates, the trainees' postprogram employment rates substantially exceeded those of the control group members. As shown by Figure 1, the trainees' and controls' preprogram employment rates were essentially identical and were declining during the two years prior to the baseline. After the baseline, the employment rates of the two groups diverged as the trainees entered NSW jobs and the controls sought regular unsubsidized employment. 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 26th month, or more than a year after the typical trainee had left the program, the employment rates of the trainees exceeded those of the control group by 9 percentage points. Therefore, the

---

4 See the Appendix for more details about the NSW sample used in this study, and Table A, which provides the means of the trainees' and controls' demographic and pre-baseline employment characteristics. For an in-depth discussion of the program and its costs, see Hollister et al. (1984).

5 Those unfamiliar with training program data may be surprised by the growth in the controls employment rates after the baseline. However, this pattern is a consistent feature of the data used in training evaluations and reflects the program's eligibility criteria (Ashenfelter (1978)).
experimental evaluation shows that at least in the short run, NSW substantially improved the employment prospects of AFDC participants.

The NSW demonstration 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. To begin our analysis of these effects of training, we examine the Kaplan-Meier survivor functions for the treatments’ and controls’ employment and unemployment spells in Table 1. The first two columns of the table indicate that 65 percent of the trainees’ employment spells lasted six or more months compared with only 57.3 percent of the controls’ spells. When we follow standard practice and compare the experience of treatments and controls in fresh unemployment spells in columns three and five of Table I, we see that 73 percent of the treatments are still in an unemployment spell after a duration of 6 months compared to only 61.3 percent of the controls. Thus training appears to be a mixed blessing since it increases the length of both employment and unemployment spells.

Unfortunately, as previously noted, such a simple analysis of the treatments’ and controls’ employment histories may be misleading. First, the possibility that the treatments and controls faced different demand conditions is particularly

---

**FIGURE 1.**—Employment rates of AFDC women in the NSW Demonstration.
TABLE I

EMPIRICAL SURVIVOR FUNCTIONS
(Proportion Remaining Employed or Unemployed)

<table>
<thead>
<tr>
<th>Months</th>
<th>Treatments (1)</th>
<th>Controls (2)</th>
<th>Treatments (3)</th>
<th>Controls: All Spells (4)</th>
<th>Controls: Fresh Spells (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>0.968 (0.013)</td>
<td>0.929 (0.018)</td>
<td>0.955 (0.013)</td>
<td>0.949 (0.011)</td>
<td>0.929 (0.023)</td>
</tr>
<tr>
<td>1</td>
<td>0.929 (0.019)</td>
<td>0.848 (0.026)</td>
<td>0.910 (0.018)</td>
<td>0.914 (0.015)</td>
<td>0.895 (0.028)</td>
</tr>
<tr>
<td>2</td>
<td>0.839 (0.027)</td>
<td>0.761 (0.030)</td>
<td>0.864 (0.021)</td>
<td>0.843 (0.019)</td>
<td>0.791 (0.039)</td>
</tr>
<tr>
<td>3</td>
<td>0.787 (0.031)</td>
<td>0.687 (0.033)</td>
<td>0.817 (0.024)</td>
<td>0.807 (0.021)</td>
<td>0.756 (0.039)</td>
</tr>
<tr>
<td>4</td>
<td>0.733 (0.033)</td>
<td>0.648 (0.034)</td>
<td>0.778 (0.025)</td>
<td>0.781 (0.022)</td>
<td>0.728 (0.039)</td>
</tr>
<tr>
<td>5</td>
<td>0.670 (0.034)</td>
<td>0.603 (0.034)</td>
<td>0.746 (0.026)</td>
<td>0.756 (0.022)</td>
<td>0.672 (0.041)</td>
</tr>
<tr>
<td>6</td>
<td>0.650 (0.035)</td>
<td>0.573 (0.035)</td>
<td>0.730 (0.027)</td>
<td>0.725 (0.023)</td>
<td>0.613 (0.043)</td>
</tr>
</tbody>
</table>

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.

pertinent in the present case because participants entered NSW as the economy was recovering from the 1974–75 recession. As a result, the controls encountered significantly worse labor market conditions during the portions of their spells that occurred while the treatment group received training.7

More importantly, as also observed in Section 2, an experimental design ensures only that the entire sample of treatments and controls are random draws from the same population. It does not ensure, for example, that the subsamples of treatments and controls experiencing employment spells are drawn randomly from the same population. To explore this possibility, we present in Table II the mean characteristics of treatments and controls experiencing different types of spells. As shown by columns one and two, these figures suggest that treatments experiencing employment spells are younger, less skilled, less likely to have ever been married, and have fewer weeks of work in the previous two years. Because these characteristics usually are associated with shorter employment durations, the NSW demonstration may have had an even larger impact on these durations than that suggested by the Kaplan-Meier estimates in Table I.

A similar problem arises in the fresh unemployment spells. As shown by columns three and five of Table II, the treatments and controls experiencing

---

7 As noted in Section 2, we could avoid this problem (but encounter others) by using only the post-training data for treatments and controls. We do not focus on this problem here since standard parametric models should be able to deal with it.
### TABLE II

**INDIVIDUAL AND SPELL CHARACTERISTICS**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Employment Treatments (1)</th>
<th>Employment Controls (2)</th>
<th>Unemployment Treatments (3)</th>
<th>Unemployment Controls: All Spells (4)</th>
<th>Unemployment Controls: Fresh Spells (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>33.77</td>
<td>34.73</td>
<td>33.21</td>
<td>34.98</td>
<td>34.36</td>
</tr>
<tr>
<td></td>
<td>(.60)</td>
<td>(.63)</td>
<td>(.51)</td>
<td>(.45)</td>
<td>(.72)</td>
</tr>
<tr>
<td>Schooling</td>
<td>10.42</td>
<td>10.55</td>
<td>10.18</td>
<td>10.11</td>
<td>10.50</td>
</tr>
<tr>
<td></td>
<td>(.14)</td>
<td>(.17)</td>
<td>(.13)</td>
<td>(.13)</td>
<td>(.20)</td>
</tr>
<tr>
<td>H.S. Dropout</td>
<td>.62</td>
<td>.60</td>
<td>.71</td>
<td>.71</td>
<td>.63</td>
</tr>
<tr>
<td></td>
<td>(.04)</td>
<td>(.04)</td>
<td>(.03)</td>
<td>(.03)</td>
<td>(.05)</td>
</tr>
<tr>
<td>Kids under 18</td>
<td>2.29</td>
<td>2.41</td>
<td>2.26</td>
<td>2.30</td>
<td>2.40</td>
</tr>
<tr>
<td></td>
<td>(.10)</td>
<td>(.12)</td>
<td>(.09)</td>
<td>(.08)</td>
<td>(.15)</td>
</tr>
<tr>
<td>Never Married</td>
<td>.38</td>
<td>.30</td>
<td>.39</td>
<td>.33</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td>(.04)</td>
<td>(.04)</td>
<td>(.03)</td>
<td>(.03)</td>
<td>(.05)</td>
</tr>
<tr>
<td>Proportion Black</td>
<td>.83</td>
<td>.78</td>
<td>.86</td>
<td>.82</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>(.03)</td>
<td>(.04)</td>
<td>(.02)</td>
<td>(.02)</td>
<td>(.04)</td>
</tr>
<tr>
<td>Prior Experience</td>
<td>2.46</td>
<td>4.09</td>
<td>2.82</td>
<td>2.91</td>
<td>5.04</td>
</tr>
<tr>
<td></td>
<td>(.63)</td>
<td>(.65)</td>
<td>(.40)</td>
<td>(.42)</td>
<td>(.87)</td>
</tr>
<tr>
<td>Proportion Never Employed</td>
<td>.17</td>
<td>.13</td>
<td>.15</td>
<td>.18</td>
<td>.12</td>
</tr>
<tr>
<td>Number of Women</td>
<td>149</td>
<td>138</td>
<td>222</td>
<td>266</td>
<td>92</td>
</tr>
<tr>
<td>Number of Spells</td>
<td>185</td>
<td>198</td>
<td>269</td>
<td>374</td>
<td>126</td>
</tr>
</tbody>
</table>

Notes: 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. The statistics in column 5 include only spells that begin after the baseline. Prior experience is measured as the number of weeks worked during the two years prior to the baseline. See Appendix Table A for the means and standard errors of the full samples of treatments and controls. The numbers in parentheses are the standard errors.

Fresh unemployment spells have quite different characteristics. The controls are significantly better educated than the treatments. Further, the controls had substantially more work experience during the previous two years. The latter difference in prior experience is quite important because it results from both observed and unobserved differences between treatments’ and controls’ characteristics.

The differing employment dynamics of the treatments’ and controls’ employment histories leads to this sample selection problem in fresh unemployment spells. Approximately 70 percent of the trainees leave NSW for a fresh unemployment spell. By contrast, 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. Not surprisingly, only one-third of the controls reach a fresh unemployment spell during the sampling period. Hence, the standard approach in event history studies of using only fresh spells leads to a sample of above-average controls.

We potentially could eliminate this sample selection problem for the controls by treating time spent after the baseline in an interrupted unemployment spell as equivalent to time spent in a fresh unemployment spell. Indeed, we see from
comparing columns three and four of Table I that when we make this assumption, 73 percent of both the trainees’ and the controls’ unemployment spells lasted at least 6 months. Moreover, a comparison of columns three and four of Table II indicates no substantial differences between the treatments and this expanded sample of controls. However, data on remaining time in interrupted unemployment spells are comparable to data from fresh unemployment spells only in the absence of duration dependence and our empirical work below reveals substantial evidence of duration dependence.

We could allow for duration dependence in our treatment of an interrupted spell by conditioning on the prebaseline duration in the spell. Unfortunately, this approach is inappropriate in the presence of unobserved heterogeneity. In general, it will contaminate the experimental design, because the controls’ heterogeneity distribution is implicitly conditioned on the prebaseline duration (Heckman and Singer (1984a, p. 108, footnote 20)). By contrast, the trainees’ interrupted spells end when they are assigned into the treatment group at the baseline, and as a result their heterogeneity distribution is not conditioned on the prebaseline duration. This problem is particularly important because it causes the treatments and controls to have different heterogeneity distributions in a case in which the benefit of random assignment is that it insured that these distributions were the same. As a practical matter, conditioning on the start date for the interrupted spells leads to a survivor function for the controls that is very similar to that based only on the fresh spells. The reason for this result is that we usually only observe controls early in their unemployment spells when they are in the midst of a fresh spell. By contrast, when we observe the controls at the baseline, they usually have already spent a substantial amount of time in their interrupted unemployment spell.

The problems associated with using only the controls’ fresh spells may be compounded by the selection process for trainees. As noted above, most trainees become unemployed when they leave NSW, but some move directly into a regular job. Further, some of these trainees do not experience a subsequent spell of unemployment during the sample period. As shown in columns one and three of Table II, trainees with employment spells are more skilled than those with unemployment spells. Consequently, the sample of trainees with unemployment spells excludes women with “above average” characteristics. Therefore, using only fresh spells to estimate the effect of training on unemployment durations causes us to compare above-average controls to below-average treatments.

As noted in the previous section, standard models of transition data can mitigate some of the foregoing problems by conditioning on demand variables and on observed characteristics, and by allowing for unobserved heterogeneity

---

8 For the sake of simplicity, we abstracted from this complication in our example in Section 2. However, the addition of sorting on the part of the treatments does not eliminate the selection problems discussed there.

9 The difference in prior experience between the treatments and controls in columns one and two indicates that there also may be selection bias in employment spells. In our empirical work, we do not find any evidence that this bias is large.
that is uncorrelated with these observed characteristics. However, because these models explicitly assume that the unobserved heterogeneity is uncorrelated with the observed variables, they cannot account for the effects of unobserved variables that are correlated with a person’s training status. In order 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) in the sample experiencing fresh unemployment spells, unobserved heterogeneity is uncorrelated with a woman’s training status. We now turn to a statistical model that avoids both of these assumptions.\textsuperscript{10}

4. ECONOMETRIC MODEL

There are several different ways to use the NSW data to construct a likelihood function for the treatments’ and controls’ employment histories. One approach utilizes both the postbaseline data and the two years of available prebaseline data. Unfortunately, this strategy leads to an extremely intricate likelihood function, even for simple specifications of the hazard functions, because of the NSW’s complex eligibility criteria. The program 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.

A second approach is to use the exact likelihood for the post baseline data without conditioning on the starting date of the interrupted spells. Unfortunately, this likelihood function also is extremely complicated.\textsuperscript{11} A third approach also utilizes only the postbaseline data, while conditioning on the starting date of the interrupted spells for the controls. As noted in the previous section, this approach will produce inconsistent estimates in the presence of unobserved heterogeneity. In light of these drawbacks, we follow Heckman’s and Singer’s (1984a) suggestion and define a separate hazard and heterogeneity term for the interrupted spells.

4.1 Contribution of the Controls’ Employment Histories

To facilitate our development of the likelihood function, we begin by describing the employment history for a hypothetical member of the control group. As shown by Figure 2, when a woman volunteers for training at the baseline (or experimental time 0), she has been unemployed for $T$ periods. After the

\textsuperscript{10} Standard models also cannot deal with the contamination to random assignment that arises from conditioning (on the prebaseline duration) in the controls’ (i) hazard for the interrupted spells and (ii) heterogeneity distribution. A referee has suggested that it may be possible to deal with this problem by conditioning both the treatments’ and controls’ heterogeneity distributions on prebaseline duration. This suggestion is beyond the scope of our paper and we have not explored it.

\textsuperscript{11} See Ham and LaLonde (1991) for derivations of both of these likelihood functions. One might expect that using the prebaseline data would alleviate our initial conditions problem. But it simply moves this problem back two years and yet still requires us to model the eligibility rules that volunteers had to satisfy in order to be admitted into the experimental sample.
program administrators randomly assign her to the control group, she remains in that spell for an additional $t_r$ periods. She next experiences an employment spell lasting $t_{e1}$ periods followed by a fresh unemployment spell lasting $t_u$ periods. Finally, she begins another employment spell which is in progress at the end of the sampling frame and thus is censored after $t_{e2}$ periods.

To form the likelihood function for this control’s employment history, we must define the transition rates out of each of the types of spells depicted in Figure 2. In what follows we work in discrete time and define these transitions rates for person $i$ as follows:\(^{12}\)

\[
\lambda_{ji}(t|\theta_{ji}) = \left(1 + \exp(-y_{ji}(t))\right)^{-1},
\]

where

\[
y_{ji}(t) = \beta_j X_i(t + t^*) + \gamma_j D_j + \delta_{ij1}\log(t) + \delta_{ij2}\log(t)^2 + \theta_{ji}.
\]

In (1) the subscript $j$ denotes the type of spell. Accordingly, $j = r$ denotes an interrupted unemployment spell; $j = u$ denotes a fresh unemployment spell; and $j = e$ denotes a fresh employment spell. The spell begins at calendar time $t^*$ and $t$ is its current duration. The vector $X_i$ includes both personal characteristics and demand variables. Among the personal characteristics are age, years of schooling, whether or not the woman dropped out of high school, the number of children less than 18, race, and marital status.\(^{13}\) The demand variables are monthly nonagricultural employment and the number of persons receiving unemployment benefits. We measure both demand variables as log deviations from SMSA means. The dummy variable $D_i$ equals 1 when a woman belongs to the treatment group and 0 otherwise. To capture the effects of duration

\(^{12}\) The hazard in (1a) is conditioned on the $X_i(\cdot)$ variables, which could in principle contain the entire history up to this period. For notational ease we make this conditioning implicit and condition explicitly only on $\theta$.

\(^{13}\) 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.
dependence, we used log duration and its square. Finally, we represent the unobserved characteristics by a scaler random variable $\theta_{ij}$. Because we leave the mean of $\theta_{ij}$ unrestricted, the parameter vector $\beta_j$ does not include an intercept. In Section 4.3, we discuss our different specifications for the joint distribution of these heterogeneity terms across different types of spells.

Using the foregoing transition rates, we can now define the contribution to the likelihood for the hypothetical control depicted in Figure 2. (In what follows we drop the $i$ subscript.) The contribution of the fresh employment and unemployment spells is straightforward (Flinn and Heckman (1983)). The contribution of the first employment spell and the fresh unemployment spell in Figure 2 conditional on $\theta_j$ is given by

\begin{equation}
(2a) \quad f(t_{el}\mid \theta_e) = \lambda_e(t_{el}\mid \theta_e) \prod_{k=1}^{t_{el}-1} (1 - \lambda_e(k\mid \theta_e))
\end{equation}

and

\begin{equation}
(2b) \quad f(t_u\mid \theta_u) = \lambda_u(t_u\mid \theta_u) \prod_{k=1}^{t_u-1} (1 - \lambda_u(k\mid \theta_u)).
\end{equation}

(Unless otherwise noted, all contributions to the likelihood are conditional on the realization of the unobserved heterogeneity.) The contribution of the right-censored employment spell in Figure 2 is given by

\begin{equation}
(3) \quad S(t_{e2}\mid \theta_e) = \prod_{k=1}^{t_{e2}} (1 - \lambda(k\mid \theta_e)).
\end{equation}

As indicated above, modeling the exact contribution of the unemployment spell in progress at the baseline is very complicated and we follow Heckman's and Singer's suggestion and let this spell have its own hazard and heterogeneity term. Therefore, the contribution for the interrupted spell in Figure 2 is given by

\begin{equation}
(4) \quad f_r(t_r\mid \theta_r) = \lambda_r(t_r\mid \theta_r) \prod_{k=1}^{t_r-1} (1 - \lambda_r(k\mid \theta_r)).
\end{equation}

Combining the information on the hypothetical control's interrupted and fresh spells, we arrive at the overall unconditional contribution to the likelihood

---

14 We also controlled for higher order polynomials of log duration, but these terms were never significant. The quadratic term was significant only for the transition rates from fresh employment spells. Because the quadratic term was not significant for the transition rates from interrupted or fresh unemployment spells, we dropped it for these transition rates. We also used time dummy variables instead of log duration and log duration squared to capture the effects of duration dependence (Ham and Rea (1987), Meyer (1990)). This alternative specification had no effect on any of the estimated coefficients, including that for training status. Finally, a referee has noted that it may be advantageous to make an adjustment when using a polynomial in log duration in a continuous time model (see Lancaster, Imbens, and Dolton (1987)).

15 We always assume that $\theta_{ij}$ is constant across spells of the same type.
function given by

\[ \mathcal{L}(t_r, t_{e1}, t_u, t_{e2}) = \int f_r(t_r|\theta_r) f_e(t_{e1}|\theta_e) f_u(t_u|\theta_u) \]
\[ \times S_e(t_{e2}|\theta_e) dG_1(\theta_r, \theta_e, \theta_u) \]

where \( G_1(\cdot) \) is the distribution function for \( \theta_r, \theta_e, \) and \( \theta_u \).

Of course there are other labor market histories than that depicted in Figure 2, but their contribution to the likelihood is analogous to (5). However, because such a large percentage of controls never left their interrupted spell during the sample period, it is worth describing their contribution to the likelihood function. The unconditional contribution of such an employment history is given by

\[ \mathcal{L}(T) = \int S_r(T|\theta_r) dG_r(\theta_r) \]

where \( T \) is the length of the postbaseline sample period,

\[ S_r(T|\theta_r) = \prod_{k=1}^{T} (1 - \lambda_r(k|\theta_r)) \]

and \( G_r(\theta_r) \) is the marginal distribution function of \( \theta_r \).

4.2 Contribution of the Treatments' Employment Histories

We now turn to the treatment group's contribution to the likelihood function. In contrast to a control's interrupted spell, a treatment's unemployment spell ends when she is randomly assigned into training. As shown for a hypothetical treatment in Figure 3, we find that after the program administrators randomly assign her to the treatment group, she remains in the program until training ends after \( S^* \) periods. She then experiences a fresh unemployment spell lasting \( t_u \) periods. At the end of this spell she begins an employment spell that is still in progress at the end of the sampling frame and thus is censored at \( t_e \) periods.

![Figure 3.—Employment history for a hypothetical treatment who completes training.](image-url)
The treatments usually were eligible for up to 12 months of training, but approximately one half dropped out of the program early to go to a regular job or to unemployment. 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 use a multiple exit framework to describe the transition into employment or unemployment after \( s \) periods of training. Accordingly, we write the conditional probability of dropping out of training into employment in week \( t_{se} \) as

\[
\lambda_{se}(t_{se}|\theta_{se}) = (1 + \exp(-y_{se}(t_{se})))^{-1},
\]

where

\[
y_{se}(t_{se}) = \beta_{se} X(t_{se} + t^*) + h_{se}(t_{se}) + \theta_{se},
\]

and where the baseline for this individual occurs at calendar time \( t^* \) and \( h_{se}(t_{se}) \) is a polynomial in log duration. Likewise we assume that the conditional probability of dropping out of training to unemployment in week \( t_{su} \) is given by

\[
\lambda_{su}(t_{su}|\theta_{su}) = (1 + \exp(-y_{su}(t_{su})))^{-1},
\]

where

\[
y_{su}(t_{su}) = \beta_{su} X(t_{su} + t^*) + h_{su}(t_{su}) + \theta_{su},
\]

and where \( h_{su}(t_{su}) \) also is a polynomial in log duration.\(^{16}\)

The individual in Figure 3 completes training and does not drop out early. Thus the training period contributes

\[
S_s(S^*|\theta_{su}, \theta_{se}) = \prod_{k=1}^{S_s} (1 - \lambda_{su}(k|\theta_{su}))(1 - \lambda_{se}(k|\theta_{se})).
\]

The contributions of the treatments’ subsequent fresh employment and unemployment spells are straightforward and analogous to those of the controls’ spells. Thus the unconditional contribution is given by

\[
\mathcal{Z}(S^*, t_u, t_e) = \int S_s(S^*|\theta_{su}, \theta_{se}) f_u(t_u|\theta_u) S_e(t_e|\theta_e) dG_2(\theta_{su}, \theta_{se}, \theta_u, \theta_e)
\]

where \( G_2(\theta_{su}, \theta_{se}, \theta_u, \theta_e) \) is the joint distribution function for \( \theta_{su}, \theta_{se}, \theta_u, \theta_e \).

To calculate the contribution for a trainee who drops out of the program early, consider the employment history shown in Figure 4. This woman drops out of training after \( t_{se} \) periods to begin a regular job. This employment spell lasts \( t_e \) periods, and is followed by an unemployment spell that is censored at the end of the sampling frame after \( t_u \) weeks. The contribution of this woman’s training

\(^{16}\) The data indicated the need for \( h_{se}(t_{se}) \) to be a second order polynomial in log duration and \( h_{su}(t_{su}) \) to be a sixth order polynomial in log duration.
SAMPLE SELECTION AND INITIAL CONDITIONS

191

state
Training
Employment
T
Unemployment

baseline=0
S*
T

FIGURE 4.—Employment history for a hypothetical treatment who drops out of training.

spell is given by

\[
 f_{se}(t_{se}|\theta_{se}, \theta_{su}) = \lambda_{se}(t_{se}|\theta_{se}) \prod_{k=1}^{t_{se}-1} (1 - \lambda_{se}(k|\theta_{se}))(1 - \lambda_{su}(k|\theta_{su})).
\]

Therefore, the unconditional contribution of the employment history is given by

\[
 \mathcal{Z}(t_{se}, t_e, t_u) = \int f_{se}(t_{se}|\theta_{se}, \theta_{su}) f_e(t_e|\theta_e) S(t_u|\theta_u) dG_2(\theta_{se}, \theta_{su}, \theta_e, \theta_u).
\]

4.3 Specification of the Heterogeneity Distributions

To estimate the effect of training on employment and unemployment duration, we maximize the likelihood formed by combining the contributions of the treatments’ and controls’ spells (Ridder and Verbakel (1984)). The resulting likelihood function is relatively complex and computationally demanding as it depends on five different parameter vectors \( (\beta_r, \beta_u, \beta_e, \beta_{su}, \beta_{se}) \) and five different heterogeneity terms \( (\theta_r, \theta_u, \theta_e, \theta_{su}, \theta_{se}) \).

We estimate this model under the following assumptions about the unobserved heterogeneity:

(A1) There is no unobserved heterogeneity, and thus

\[
 \theta_j = \alpha_j, \quad j = se, su, r, u, e,
\]

where \( \alpha_j \) is a fixed intercept.

(A2) Unobserved heterogeneity is independent across spells and is drawn from a two-point distribution

\[
 \theta_j = \begin{cases} 
 \theta_{j1} & \text{with probability } P_j, \\
 \theta_{j2} & \text{with probability } 1 - P_j,
\end{cases}
\]

where \( j = r, u, e \).

\[17\] In all but (A5) below, we assume that \( \theta_{se} \) and \( \theta_{su} \) follow a one factor structure with one intercept normalized to zero and one loading factor normalized to one.
(A3) The unobserved heterogeneity components \((\theta_u, \theta_e)\), \(\theta_r\) are independent of each other and of \(\theta_u\) and \(\theta_e\). Further \(\theta_u\) and \(\theta_e\) follow the one-factor structure

\[ \theta_j = \alpha_j + c_j \theta^* \]

where \(j = u, r\) and \(\theta^*\) is drawn from a two-point distribution. (In this specification and in (A4) and (A5) below, we normalize \(\alpha_u = 0\) and \(c_u = 1\).)

(A4) The unobserved heterogeneity terms \((\theta_{su}, \theta_{se})\) are independent of \(\theta_u\), \(\theta_e\), and \(\theta_r\). Further \(\theta_u\), \(\theta_e\), and \(\theta_r\) follow the one-factor structure given by (18).

(A5) All heterogeneity terms follow the one-factor structure given by (18).\(^{18}\)

Assumptions (A1) and (A2) imply that there will be no selection bias from using only fresh employment and unemployment spells in estimation. Our discussion of the survivor functions in Table I and of the means of the treatments' and controls' characteristics in Table II suggests that these assumptions are inappropriate within the current parametric framework.

The Assumptions (A3) through (A5) allow us to address the sample selection problem. Assumption (A3) focuses on the idea that the crucial correlation is between the controls' interrupted spells and fresh unemployment spells. Assumption (A4) allows for the possibility that there is additional sample selection bias arising from treating the unobservables in the employment spells as independent from unobservables in the interrupted and fresh unemployment spells. For example, controls who leave an interrupted spell and enter an employment spell may have above average unobserved characteristics. Finally, Assumption (A5) addresses the possibility that there is sample selection among the treatments depending on how they leave training. For example, we would expect treatments who drop out of training for employment to have above-average unobserved characteristics.

In our empirical work, we obtain estimates for two different specifications of the interrupted spell hazard, \(\lambda_r\). In the first case, we do not condition on the time spent in the interrupted spell before the baseline. In the second case, we do condition on the time spent in the spell prior to the baseline. Specifically, in this later case, we measured duration in the interrupted spells from the beginning of those spells and not from the baseline. We noted above that if we use the same hazard and heterogeneity distribution for the fresh and interrupted unemployment spells (with duration in the latter spells measured from the beginning of the spell), we will contaminate the experimental design. However, when we allow the interrupted spells to have their own hazard,\(^{18}\)

\(^{18}\) The project was quite computationally demanding and thus we generally stayed with the assumption that \(\theta^*\) was drawn from a two-point distribution. However, we did try to add a third point of support for one of our more complex models. 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 one of our most complicated models, we maintained the assumption that \(\theta^*\) was drawn from a two-point distribution in our other specifications.
accounting for time in this spell prior to the baseline simply helps us obtain a better approximation to the interrupted spell hazard and a greater degree of (empirical) identification by exploiting the variation in this variable among the control group members. Therefore, our alternative specifications for $\lambda_r$ should provide a check on the robustness of our results.

In an earlier version of the paper (Ham and LaLonde (1991)), we estimated the model under Assumptions (A3) through (A5) for the case in which we do not condition on prebaseline duration and $\alpha_j$ is restricted to zero in (18) for all $j$. This restriction places a proportionality restriction on the intercept terms for a woman of a given type. For example, this restriction constrains the constant in the interrupted hazard to be $c_r$ times the intercept in the fresh unemployment spell and to be $c_e/c_r$ times the intercept in the fresh employment spell. Our estimates under Assumptions (A3), (A4), and (A5) generally produced quite similar estimates for the cases in which (i) we condition on prebaseline duration in the interrupted spells; (ii) we do not condition on prebaseline duration and do not restrict $\alpha_j = 0$ in (18); and (iii) we do not condition on prebaseline duration and restrict $\alpha_j = 0$ in (18). Accordingly, we present results for the case in which we condition on prebaseline duration in the interrupted spell. The only real difference among the estimates occurred when we estimated the full one-factor model (A5) and did not condition on prebaseline duration in the interrupted spells nor imposed the restriction that $\alpha_j = 0$ in (18). In this case we essentially lost identification.\footnote{We found three distinct optima. One looked very similar to the estimates presented below in columns three and five of Table IV. For the other two optima, the estimates implied that one of the hazards was zero (for one of the mass points of distribution) in the sense of being smaller than $e^{-66}$; moreover, in each case the maximization algorithm appeared to be exploiting the fact that as it changed the heterogeneity distribution, the computer was treating a hazard for one heterogeneity term as equal to zero. One of these optima yielded training effects quite similar to those in Table IV. However, the other optima yielded a large positive training effect in the employment hazard and a large negative training coefficient in the unemployment hazard.} Although it is hard to generalize from one data set, these results suggest that the prebaseline duration data provide important identifying information. Further, in the absence of such information, it may be necessary to impose additional restrictions such as placing a proportionality restriction on the intercepts.

5. ESTIMATES OF THE TRAINING EFFECT

We first present estimates of the training effects assuming no heterogeneity (A1) or independent heterogeneity (A2) among spells. These assumptions imply that we can obtain consistent estimates of the training effect using only the fresh spells. Next, we show how incorporating information on the controls’ interrupted spells affects our estimates of the fresh unemployment hazard using Assumption (A3). Then we examine the effect of including information from the interrupted spells on our estimates of both the fresh employment and unemployment hazards using Assumption (A4). Finally, we include the treatments’ training
TABLE III
ESTIMATED TRAINING EFFECTS USING FRESH SPELLS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Employment</th>
<th>Unemployment</th>
<th>Fresh Spells Only</th>
<th>Fresh and Interrupted Spells Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Training Status</td>
<td>-.394</td>
<td>-.425</td>
<td>-.382</td>
<td>-.374</td>
</tr>
<tr>
<td></td>
<td>(.155)</td>
<td>(.180)</td>
<td>(.166)</td>
<td>(.208)</td>
</tr>
<tr>
<td>Log Duration</td>
<td>.212</td>
<td>.155</td>
<td>-.453</td>
<td>-.299</td>
</tr>
<tr>
<td></td>
<td>(.246)</td>
<td>(.258)</td>
<td>(.075)</td>
<td>(.119)</td>
</tr>
<tr>
<td>Log Duration Squared</td>
<td>-.168</td>
<td>-.113</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(.069)</td>
<td>(.085)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Controls for Unobserved Heterogeneity</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-852.7</td>
<td>-851.6</td>
<td>-768.1</td>
<td>-765.5</td>
</tr>
<tr>
<td></td>
<td>(.069)</td>
<td>(.085)</td>
<td>(.141)</td>
<td>(.1411)</td>
</tr>
</tbody>
</table>

Notes: All models also include controls for age, years of schooling, a woman’s high school dropout status, number of children under 18, marital status, race, and a woman’s SMSA’s establishment employment and unemployment insurance recipients. Local employment and unemployment are defined as the log deviations from site means. Appendix Table B presents the estimated coefficients for these variables. The standard errors are in parentheses.

As was suggested by the empirical survivor functions in Table I, the estimated training effects based on the fresh spells imply that training lowered both the probabilities of leaving employment and unemployment. Column one of Table III presents the coefficient (and standard error) of the training dummy variable and the duration terms from the employment hazard when $\theta_e$ is constant, while column two presents the same estimates when this heterogeneity term is assumed to come from a discrete distribution with two points of support. (The full set of parameter estimates is contained in Appendix Table B.) The estimate of $-.394$ implies that training increased expected employment duration by approximately 11 months. In columns three and four of Table IV, we present the corresponding estimates for the unemployment hazard. The estimate of $-.382$ implies that training increased expected unemployment duration by approximately 40 months.20

Although the estimated effect of training on unemployment durations is not credible, the other estimated coefficients are economically plausible. As shown.

20 We calculated the estimated differences in trainees’ and controls’ expected durations using the parameters from the estimated hazard functions which ignore unobserved heterogeneity.
SAMPLE SELECTION AND INITIAL CONDITIONS

TABLE IV
ESTIMATES OF THE TRAINING EFFECT BASED ON ALTERNATIVE HETEROGENEITY ASSUMPTIONS

<table>
<thead>
<tr>
<th></th>
<th>Fresh Unemployment Hazard</th>
<th>Employment Hazard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)^a</td>
<td>(2)^b</td>
</tr>
<tr>
<td>Training</td>
<td>-0.073</td>
<td>-0.062</td>
</tr>
<tr>
<td>Status (.209)</td>
<td>(.209)</td>
<td>(.207)</td>
</tr>
<tr>
<td>Log Duration</td>
<td>-0.357</td>
<td>-0.370</td>
</tr>
<tr>
<td>Log Duration Squared</td>
<td>(.070)</td>
<td>(.095)</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-1401.2</td>
<td>-2251.9</td>
</tr>
<tr>
<td>Log Likelihood Squared</td>
<td>(.069)</td>
<td>(.069)</td>
</tr>
<tr>
<td>No Heterogeneity</td>
<td>-1410.2</td>
<td>-2262.8</td>
</tr>
</tbody>
</table>

Notes: See notes to Table III; estimates in the table correspond to the assumptions described below and in the text. 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).

(A3), where the heterogeneity terms for r and u follow the one-factor structure given by (18). Under (A3) the corresponding estimates for the employment hazard are from the model with independent heterogeneity, reported in column 2 of Table III.

(A4), where the heterogeneity terms for r, u, and e follow the one-factor structure given by (18).

(A5), where all heterogeneity terms follow the one-factor structure given by (18).

by Appendix Table B, individuals have longer unemployment spells if they are older, less educated, or African-American. In addition, adverse local labor market conditions also seem to increase the duration of unemployment. By contrast, the only personal characteristic besides training status that significantly affects the employment hazards is high school dropout status. Women who are high school dropouts experienced significantly shorter employment spells. Finally, we find substantial duration dependence both in the unemployment and employment hazards. The quadratic term in the employment hazard reflects the tendency in the data for the hazard to rise during the first few weeks in a new job before declining in subsequent weeks.

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 training coefficient merely reflects the selection problem discussed above. Moreover, we would not expect the problem to be resolved by accounting for observed characteristics and unobserved heterogeneity (in column four), because the standard approach to unobserved heterogeneity assumes that the heterogeneity terms are independent of the explanatory variables, including training status. We have argued above that our sample selection problem contaminates the experimental design and thereby causes a woman’s training status to be correlated with the heterogeneity term in the fresh unemployment hazard.

An informal way of avoiding this selection problem is to include the controls’ interrupted spells in the analysis and to assume that the fresh and interrupted
spells have the same hazard function. Of course, this procedure is valid only under the strong assumption of no duration dependence in the unemployment hazard. If this 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 remaining in a fresh spell—against the sample selection bias that results from excluding these interrupted spells.

As shown by column five, when we use the same hazard for both the 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, and this result is not simply caused by unobserved heterogeneity. As shown in column six, when we account for unobserved heterogeneity, 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 use the same hazard function for the interrupted and fresh unemployment spells.

5.2 Results from Models Allowing for Sample Selection

The findings in Table III underscore the potential importance of sample selection and motivate consideration of the more complicated statistical framework developed in this paper. We first obtain estimates based on Assumption (A3) where the heterogeneity terms $\theta_u$ and $\theta_r$ follow the one-factor structure given by (18) and are independent of $\theta_s$, $\theta_{se}$, and $\theta_{su}$. As shown in column one of Table IV, now training has almost no effect on the transition rate out of a fresh unemployment spell. This result contrasts with the corresponding estimate in column four of Table III, which indicated that training substantially increased unemployment durations. Further, this result is essentially identical to the one obtained when we forced the interrupted and fresh spells to have the same hazard function (in column six of Table III). This finding suggests that the bias that results when we follow standard practice and discard the interrupted spells

\footnote{For the estimates presented in Table III, we measure the duration in the spells as time spent unemployed since the baseline. Alternatively, if we use the total duration of these spells, which includes time spent unemployed prior to the baseline, the estimated training effect is similar to that reported for fresh spells in column three. If there were no unobserved heterogeneity, this alternative approach would be appropriate. But, as discussed in Section 3, in the presence of such heterogeneity, conditioning on prebaseline duration contaminates the experimental design.}

\footnote{In what follows, we allow the interrupted and fresh unemployment spells to have different hazard functions.}

\footnote{Given Heckman and Singer's (1984b) results suggesting that our approach is likely to do a poor job of recovering the heterogeneity distribution, we do not discuss the parameter estimates for this distribution.}
is more serious than the bias that results when we treat the two types of spells as having the same hazard function.

We next impose Assumption (A4) where the heterogeneity terms $\theta_e$, $\theta_u$, and $\theta_r$ follow the one factor structure (18) and are independent of the terms associated with the training spells ($\theta_{ue}, \theta_{ur}$). As shown by column two of Table IV, when we correct for selection bias in this fashion, training again has no effect on unemployment durations. However, as shown by column four, training continues to significantly lengthen employment durations.

Finally, we account for the potential selection bias arising from the treatments’ exit from training by adopting Assumption (A5) and using the one-factor structure (18) for all the heterogeneity terms. Column three of Table IV contains the coefficient for the training dummy for the fresh spell unemployment hazard, while column five contains the estimate of the training coefficient for the employment hazard. Comparing columns two and three, we see that the training dummy in the unemployment hazard is still essentially zero. Comparing columns four and five, we see that the training coefficient in the employment hazard also is unchanged and is still quite significant. (As shown by Appendix Table B, we also find that the estimated coefficients for the other independent variables are unaffected by changing the assumed heterogeneity structure.)

Our estimates that account for sample selection indicate that training significantly increased the duration of employment spells, while it had no effect on the duration of unemployment spells. These findings make considerably more economic sense than the results based only on fresh spells, which suggest that training substantially raised unemployment duration. Our 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.

5.3 Tests of Model Specification

We also considered goodness-of-fit tests for our estimates under the assumption of independent heterogeneity (A2), which ignores selection bias and the full factor model (A5), which accounts for selection bias. We initially had hoped to test these models by comparing the predicted employment rates generated by them to the actual employment rates from a resurvey of NSW women. This survey was conducted during the Fall of 1979, which for most women was nearly one year after the end of the original sampling period. Although the resurvey sought to interview a random sample of the NSW treatments and controls, we found substantial evidence of attrition bias in the resurvey data. Not only were the nonrespondents younger than the respondents, but their employment rates during the original sampling period also were significantly different. For the controls, by the 26th month after the baseline, the nonrespondents’ employment rates were only one half as large as those of the respondents. For the treat-
TABLE V  
ACTUAL AND PREDICTED EMPLOYMENT RATES

<table>
<thead>
<tr>
<th>Months Since the Baseline</th>
<th>Controls:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Chi-Square Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Actual</td>
<td>.282</td>
<td>.289</td>
<td>.297</td>
<td>.301</td>
<td>.320</td>
<td>.289</td>
</tr>
<tr>
<td>Predicted: Model (A5)</td>
<td>.269</td>
<td>.273</td>
<td>.278</td>
<td>.284</td>
<td>.290</td>
<td>.295</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted: Model (A2)</td>
<td>.267</td>
<td>.271</td>
<td>.277</td>
<td>.283</td>
<td>.290</td>
<td>.295</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatments:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>.367</td>
<td>.364</td>
<td>.382</td>
<td>.385</td>
<td>.385</td>
<td>.382</td>
</tr>
<tr>
<td>Predicted: Model (A5)</td>
<td>.341</td>
<td>.345</td>
<td>.353</td>
<td>.359</td>
<td>.365</td>
<td>.371</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted: Model (A2)</td>
<td>.360</td>
<td>.364</td>
<td>.369</td>
<td>.373</td>
<td>.378</td>
<td>.382</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: We calculated the predicted employment rates generated by our model under the Assumptions (A5) and (A2) as defined in the text. The figures in the table are the means of each woman's predicted employment probabilities in each period. The chi-square statistic is for the joint hypothesis that the difference between the actual and predicted employment rates equal zero during months 21 through 26. The numbers in parentheses are the "p-values."

ments, the nonrespondents employment rates grew during the last nine months of the original sampling period, while the respondents employment rates remained constant during that period. Thus the resurvey data appears to contain above-average controls and below-average treatments, and thus are inappropriate for goodness-of-fit tests.

We then carried out goodness-of-fit tests for the employment rates in the last six months of the period used for estimation. Specifically, we tested the joint hypothesis that the difference between the predicted and actual employment rates for the treatments were zero in each of the six months preceding the end of the sample. We also carried out this test for the controls. We followed the approach of Heckman and Walker (1990) and used Monte Carlo simulation to calculate the employment probability for an individual over the period. As shown by Table V, we find that the full specification (A5) and the independent heterogeneity (A2) yield predicted employment rates for both treatments and controls that are nearly identical to the actual rates during the last six months of the sampling frame. Thus, we cannot distinguish between the models using goodness-of-fit statistics in the sample used for estimation. In retrospect, this result is perhaps unsurprising because both models have a relatively large

24 As in the Heckman and Walker (1990) study, our test statistics do not reflect the error in estimation of the parameters of the likelihood function because of the computational resources this would require.
number of parameters with which to fit the sample period data. In the absence of a random sample in the resurvey data, goodness-of-fit statistics do not help us distinguish between the models.

6. CONCLUSIONS

In this paper we found that NSW 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 employment 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 (1992) supports this conjecture. Using quarterly Social Security earnings data, he reports that the NSW treatments had significantly greater earnings than the controls more than seven years after the NSW 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 (1993) uses the National Longitudinal Survey to evaluate the impact of public sector training on employment and unemployment durations. His study differs from ours in two fundamental ways because he does not have an experimental design.25 First, because he does not condition the heterogeneity distribution on 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, because he must allow for individuals entering training both before and during his sampling frame, he faces a much 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.26

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

---

25 He must also aggregate across different training programs.
26 See Ridder (1986) for an evaluation of Dutch training programs with nonexperimental data. As Ridder explicitly notes, he is forced to make strong identifying assumptions because he lacks a control group and must instead rely on pre/post comparisons between unemployment durations.
do not see this as a 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. As our results indicate, this task is clearly much more manageable than Gritz's. 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.

Dept. of Economics, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.
and
Dept. of Economics, Michigan State University, East Lansing, MI 48824, U.S.A.,
and National Bureau of Economic Research, Cambridge, MA 02138, U.S.A.

Manuscript received November, 1991; final revision received February, 1995.

APPENDIX
DESCRIPTION OF NATIONAL 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,” Mathematical 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 three of the previous six months, to have received AFDC payments for thirty of the previous thirty-six 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 were 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 semimonthly intervals for the two years prior to the baseline. This information was used to calculate the respondent's number of semimonthly 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 nine-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 twenty-six months following the baseline. Some women with a twenty-seven month interval had employment data for only twenty-six months because their interview took place before the end of the month. The post-baseline employment histories in our study extend for fifty-two semimonthly periods. The sample used in the study consists of only those women with a baseline and three nine-month interviews and who satisfied the two employment related eligibility criteria for the program. Unfortunately, less than 40 percent of the sample was interviewed after twenty-seven 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 the 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—nineteen trainees and fifteen 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 three of the

<table>
<thead>
<tr>
<th>TABLE A</th>
<th>CHARACTERISTICS OF TREATMENTS AND CONTROLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatments (1)</td>
</tr>
<tr>
<td>Age</td>
<td>33.67 (.47)</td>
</tr>
<tr>
<td>Schooling</td>
<td>10.17 (.11)</td>
</tr>
<tr>
<td>H.S. Dropout</td>
<td>.70 (.03)</td>
</tr>
<tr>
<td>Number of Kids</td>
<td>2.25 (.08)</td>
</tr>
<tr>
<td>Never Married</td>
<td>.38 (.03)</td>
</tr>
<tr>
<td>Black</td>
<td>.85 (.02)</td>
</tr>
<tr>
<td>Prior Experience(^a)</td>
<td>2.59 (.34)</td>
</tr>
<tr>
<td>Never Employed(^b)</td>
<td>.16 (.02)</td>
</tr>
<tr>
<td>Number of Women</td>
<td>275</td>
</tr>
</tbody>
</table>

\(^a\) Prior experience is the number of weeks of employment in the two years prior to the baseline.

\(^b\) Never employed is a dummy variable indicating that the woman has not had a regular job since she was 16 years old.
six 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) present 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 fourteen 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

### TABLE B
**ESTIMATES OF EMPLOYMENT AND UNEMPLOYMENT HAZARD FUNCTIONS**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Employment Spells (1)</th>
<th>Unemployment Spells (2)</th>
<th>Fresh Employment Spells (3)</th>
<th>Unemployment Spells (4)</th>
<th>All Employment Spells (5)</th>
<th>All Unemployment Spells (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training status</td>
<td>-0.394</td>
<td>-0.425</td>
<td>-0.382</td>
<td>-0.374</td>
<td>-0.191</td>
<td>-0.105</td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td>(0.180)</td>
<td>(0.166)</td>
<td>(0.208)</td>
<td>(0.129)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.013</td>
<td>-0.016</td>
<td>-0.009</td>
<td>0.009</td>
<td>-0.016</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.015)</td>
<td>(0.009)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Schooling</td>
<td>0.049</td>
<td>0.056</td>
<td>0.125</td>
<td>0.149</td>
<td>0.150</td>
<td>0.169</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.062)</td>
<td>(0.063)</td>
<td>(0.078)</td>
<td>(0.068)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>H.S. Dropout</td>
<td>0.398</td>
<td>0.472</td>
<td>-0.306</td>
<td>-0.427</td>
<td>-0.335</td>
<td>-0.437</td>
</tr>
<tr>
<td></td>
<td>(0.199)</td>
<td>(0.238)</td>
<td>(0.219)</td>
<td>(0.277)</td>
<td>(0.163)</td>
<td>(0.200)</td>
</tr>
<tr>
<td>Kids under 18</td>
<td>0.010</td>
<td>0.009</td>
<td>-0.037</td>
<td>-0.024</td>
<td>0.030</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.066)</td>
<td>(0.060)</td>
<td>(0.073)</td>
<td>(0.045)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>Never Married</td>
<td>0.075</td>
<td>0.028</td>
<td>-0.148</td>
<td>-0.182</td>
<td>-0.243</td>
<td>-0.284</td>
</tr>
<tr>
<td></td>
<td>(0.167)</td>
<td>(0.205)</td>
<td>(0.182)</td>
<td>(0.238)</td>
<td>(0.134)</td>
<td>(0.167)</td>
</tr>
<tr>
<td>Black</td>
<td>0.060</td>
<td>0.065</td>
<td>-0.391</td>
<td>-0.458</td>
<td>-0.386</td>
<td>-0.469</td>
</tr>
<tr>
<td></td>
<td>(0.191)</td>
<td>(0.225)</td>
<td>(0.210)</td>
<td>(0.275)</td>
<td>(0.153)</td>
<td>(0.191)</td>
</tr>
<tr>
<td>Area Employment</td>
<td>1.78</td>
<td>1.10</td>
<td>4.60</td>
<td>3.36</td>
<td>0.531</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>(3.78)</td>
<td>(4.18)</td>
<td>(4.24)</td>
<td>(4.83)</td>
<td>(3.10)</td>
<td>(3.38)</td>
</tr>
<tr>
<td>Area Unemployment</td>
<td>0.651</td>
<td>0.669</td>
<td>-0.117</td>
<td>-0.135</td>
<td>-1.26</td>
<td>-1.19</td>
</tr>
<tr>
<td></td>
<td>(0.431)</td>
<td>(0.447)</td>
<td>(0.511)</td>
<td>(0.531)</td>
<td>(0.368)</td>
<td>(0.385)</td>
</tr>
<tr>
<td>Log Duration</td>
<td>0.212</td>
<td>0.155</td>
<td>-0.453</td>
<td>-0.299</td>
<td>-0.503</td>
<td>-0.353</td>
</tr>
<tr>
<td></td>
<td>(0.246)</td>
<td>(0.258)</td>
<td>(0.075)</td>
<td>(1.119)</td>
<td>(0.055)</td>
<td>(0.082)</td>
</tr>
<tr>
<td>Log Duration$^2$</td>
<td>-0.168</td>
<td>-0.113</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.845)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>-3.15</td>
<td>-2.78</td>
<td>-2.91</td>
<td>-2.21</td>
<td>-3.16</td>
<td>-2.34</td>
</tr>
<tr>
<td></td>
<td>(0.817)</td>
<td>(1.03)</td>
<td>(0.946)</td>
<td>(1.36)</td>
<td>(0.694)</td>
<td>(0.885)</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>—</td>
<td>-4.12</td>
<td>—</td>
<td>-3.81</td>
<td>—</td>
<td>-3.94</td>
</tr>
<tr>
<td></td>
<td>(1.19)</td>
<td></td>
<td></td>
<td>(1.36)</td>
<td>(0.862)</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>—</td>
<td>0.181</td>
<td>—</td>
<td>-1.24</td>
<td>—</td>
<td>-1.68</td>
</tr>
<tr>
<td></td>
<td>(1.20)</td>
<td></td>
<td></td>
<td>(1.50)</td>
<td>(0.836)</td>
<td></td>
</tr>
<tr>
<td>$-\log L$</td>
<td>852.7</td>
<td>851.6</td>
<td>768.1</td>
<td>765.5</td>
<td>1416.4</td>
<td>1411.9</td>
</tr>
</tbody>
</table>
### TABLE B CONT.

**ESTIMATES OF EMPLOYMENT AND UNEMPLOYMENT HAZARD FUNCTIONS**

#### B: Full Set of Estimates for Table IV

<table>
<thead>
<tr>
<th>Variable</th>
<th>Excluding Training and Employment Spells</th>
<th>Excluding Training Spells</th>
<th>Including Training Spells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unemployment Interrupted (1) Fresh (2)</td>
<td>Unemployment Interrupted (3) Fresh (4) Employment (5)</td>
<td>Unemployment Interrupted (6) Fresh (7) Employment (8)</td>
</tr>
<tr>
<td>Training status</td>
<td>— — .073 (.201) — .062 (.207) — .365 (.159)</td>
<td>— .025 (.217) — .403 (.156)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.021 (.18) .009 (.13) .021 (.19) .010 (.13) .012 (.11)</td>
<td>.020 (.17) — .0097 (.13) — .012 (.11)</td>
<td></td>
</tr>
<tr>
<td>Years of schooling</td>
<td>.200 (.088) .150 (.072) .207 (.089) .145 (.072) .057 (.053)</td>
<td>.205 (.086) .150 (.072) .055 (.052)</td>
<td></td>
</tr>
<tr>
<td>Kids under 18</td>
<td>.223 (.077) — .035 (.067) .220 (.078) — .033 (.066) .018 (.056)</td>
<td>.221 (.076) — .041 (.067) .018 (.056)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>.596 (.263) — .422 (.246) .618 (.265) — .420 (.241) .051 (.194)</td>
<td>— .603 (.257) — .442 (.205) .038 (.193)</td>
<td></td>
</tr>
<tr>
<td>Never married</td>
<td>.224 (.267) — .140 (.211) .284 (.309) — .155 (.215) .088 (.171)</td>
<td>— .222 (.265) — .101 (.212) .095 (.169)</td>
<td></td>
</tr>
<tr>
<td>H.S. dropout</td>
<td>— .595 (.304) — .516 (.264) .506 (.308) — .503 (.264) .366 (.202)</td>
<td>— .590 (.298) — .559 (.259) .373 (.201)</td>
<td></td>
</tr>
<tr>
<td>Local employment</td>
<td>— 5.76 (.554) 4.71 (.50) — 4.61 (.567) 4.50 (.447) 2.17 (3.81)</td>
<td>— 5.20 (.548) 5.27 (.450) 2.46 (.382)</td>
<td></td>
</tr>
<tr>
<td>Local unemployment</td>
<td>— 2.50 (.619) — .189 (.521) 2.44 (.617) — .198 (.520) .599 (.435)</td>
<td>— 2.42 (.614) — .206 (.517) .617 (.433)</td>
<td></td>
</tr>
<tr>
<td>Log duration</td>
<td>— .803 (.198) — .357 (.097) .816 (.197) — .370 (.095) .207 (.246)</td>
<td>— .807 (.194) — .357 (.091) .212 (.246)</td>
<td></td>
</tr>
<tr>
<td>Log duration squared</td>
<td>— — — — — 163 (.069) — — — — — 168 (.069)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant: ( a_I )</td>
<td>5.71 (3.07) 0 (3.34) 6.27 (1.07) 0 (2.81) — 2.22 (1.07) 5.27 (2.81) 0 (2.97)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta_1 )</td>
<td>— 3.76 (1.11) — 2.27 (1.05) — 3.67 (1.11) — 3.78 (1.10) — 2.11 (1.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>— 2.23 (1.06) — 3.67 (1.11) — 3.78 (1.10) — 2.11 (1.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu )</td>
<td>2.03 (.329) — 1.99 (.368) — 2.11 (.313)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loading</td>
<td>2.08 (.652) 1.0 (.732) 2.28 (.732) 1.0 (.221) .329 (.517) 1.97 (1.517) 1.0 (.173)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log ( L )</td>
<td>— 1401.2 (.652) — 2251.9 (.732) — 2994.4 (.221)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log ( L ), No heterogeneity</td>
<td>— 1410.2 (.562) — 2262.8 (.732) — 3007.9 (.221)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** See notes to Table III. The term \( \mu \) is defined as \( P_1 - e^{\mu c}/(1 + e^{\mu c}) \). In Panel B, the “loading” factors are the \( c_j \) terms in equation (18), where \( c_w \) is normalized to equal 1.
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 nine to eighteen 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 percent of the participants had left by the 13th month and 95 percent were out of the program by the 17th month.

REFERENCES


