The Impact of Being Offered and Receiving Classroom Training on the Employment Histories of Disadvantaged Women: Evidence from Experimental Data

CURTIS EBERWEIN
McGill University

JOHN C. HAM
University of Pittsburgh

and

ROBERT J. LALONDE
Michigan State University

and

NBER

First version received March 1994; final version received June 1997 (Eds.)

We address two questions using experimental data on disadvantaged women. First, what is the impact of being offered JTPA classroom training on the duration of unemployment and employment? Second, what is the effect of actually participating in this training on the length of such spells? Belonging to the treatment group shortens unemployment spells but has no effect on employment spells. Actually participating in training has a larger positive effect on the exit rate from unemployment than the effect of simply being a member of the treatment group. Ignoring the endogeneity of actual training in estimation substantially underestimates its effect.

INTRODUCTION

There is a growing consensus in the United States that public sector sponsored training programmes can improve the employment prospects of economically disadvantaged adult women (Gurone and Pauly (1991) and LaLonde (1995)). This consensus is reinforced by a recent experimental study of adult women who were offered access to classroom training administered under the U.S. Job Training Partnership Act (JTPA) (Bloom et. al. (1993)). Studies suggest that these gains usually result more from increased weeks of employment than from increased hourly wages or increased hours per week for those who are employed.

In this paper we address two questions about the impact of classroom training (CT) on the labour market histories of adult women who participated in the National JTPA Study. First, we ask whether being a member of the treatment group increased employment
rates by increasing the durations of their employment spells and/or by reducing the durations of their unemployment spells. We refer to these impacts as the effects of the "intention to treat" or the "treatment effects".

The second question we address is what are the impacts of actually receiving training on employment and unemployment durations? We refer to these as the "training effects." The second question differs from the first for at least two reasons. First, many women assigned to the treatment group are "no-shows" who never enroll in training and therefore do not receive any training services. Second, although officials prevented controls from obtaining JTPA services, they could not prevent them from obtaining similar training elsewhere in their community, and thus some members of the control group received training.

It is more difficult to estimate the training effects than the treatment effects because training status is determined by individuals' decisions and must be treated as endogenous. Fortunately, in this case treatment status no longer enters the transition rates and is an excellent predictor of training status. Because treatment status is (by design) independent of observable and unobservable characteristics, experimental data are of substantial value in identifying the training effect.

However, an experimental design does not guarantee that standard estimation approaches provide unbiased estimates of the treatment effects. This problem arises because randomization at the baseline (the point in time when volunteers are randomly assigned to the treatment and control groups) does not insure that the treatment incidence is independent of the unobservables in subsequent employment and unemployment spells. For example, when training is successful, it helps unemployed treatments become employed who have characteristics that would otherwise make them less likely to find a job. As a result, a comparison of the hazard rates of treatments and controls in the subsequent employment spell may suffer from a dynamic selection bias that arises because of the differential sorting of treatments and controls into this spell.\(^1\) Ham and LaLonde (1996) found that this bias was substantial and led to misleading policy implications when they examined adult women who participated in a work experience (WE) training programme. In this paper we extend their econometric framework to estimate the treatment effects in the JTPA data.

Our empirical results indicate that access to classroom training raises employment rates by shortening the unemployment spells of adult women who participated in JTPA-CT. Access to this training had no impact on the length of their employment spells. The effects of access to JTPA-CT are the opposite to those found by Ham and LaLonde (1996) for a WE programme; the WE programme increased employment durations but had no effect on unemployment durations. Further, unlike Ham and LaLonde (1996), we find no evidence that dynamic selection bias affects our estimates of the treatment effects in subsequent employment and unemployment spells.

We find that the actual receipt of training significantly shortens the length of unemployment spells, but does not affect the length of employment spells. (These effects are estimated relatively precisely.) Further, these training effects on unemployment durations are larger than the corresponding treatment effects. We also find that it is important to account for the endogeneity of training, because treating this variable as exogenous substantially lowers its estimated (positive) impact on labour market histories.

---

1. The intuition is similar to that described by Chesher and Lancaster (1983), who compare the correlations among observed variables and unobserved variables in samples derived from flow sampling and from stock sampling. See Section 5.
The paper proceeds as follows. In Section 2 we describe some institutional features of the JTPA programme, summarize the findings from the National JTPA Study for adult women and compare the treatments' and controls' empirical hazard rates out of employment and unemployment spells. In Section 3 we present a statistical framework for estimating treatment effects while allowing for dynamic selection bias. We then discuss our approach to estimating training effects while allowing for endogenous training in addition to this selection bias. Our empirical estimates are presented in Section 4. In Section 5 we examine why this form of dynamic selection bias was important when evaluating a WE programme but not when evaluating JTPA–CT, and discuss the implications of this difference for evaluation of other training programmes. Section 6 concludes the paper.

2. JTPA AND THE NATIONAL JTPA STUDY

2.1. The JTPA programme

Each year the U.S. government spends approximately $4 billion on JTPA programmes that provide employment and training services to economically disadvantaged persons and displaced workers. The services provided to economically disadvantaged adult women usually fall into one of several categories. We focus on vocational classroom instruction (CT), which provides training in areas such as auto mechanics, cooking, truck driving, and appliance repair, or skills useful for occupations in the clerical, health-related, and security fields. Studies of the receipt of training by disadvantaged persons indicate that programme administrators are most likely to assign adult women to programmes offering CT or basic remedial education (National Commission (1987), Smith (1995)).

JTPA has two institutional features that affect the design and interpretation of any evaluation of its impacts. First, although JTPA is a national programme, it is administered by states and localities. Often, twenty or more subcontractors provide the actual training within a locale. These subcontractors may include local community colleges, vocational schools, and private for-profit and non-profit organizations. The consequence of this administrative structure is that it is difficult to characterize JTPA training services because they vary substantially both among and within localities.

Second, JTPA programme administrators do not simply assign eligible applicants to a single service. Instead, at the time when they determine that an applicant is eligible for the programme, they also devise a training plan for the applicant. This plan often consists of a specific sequence of services. For example, administrators might recommend that less “job ready” trainees first receive basic education, next participate in vocational classroom instruction, and finally receive job search assistance. Because of these individualized training plans, when analysts evaluate the impacts of a specific component of JTPA, it is likely that the trainees have received other JTPA services besides the one under study.

2.2. The National JTPA study

The data used in this paper are from the National JTPA Study (Bloom et al. (1993)). This study used an experimental design to estimate the impact that access to JTPA services had on participants admitted into the programme during 1988 and 1989 in 16 sites across the United States. The Study implemented the experimental design after programme administrators determined an applicant’s programme eligibility and devised a training plan.

2. The appendix contains a detailed discussion of the design of the National JTPA Study and the data used in this paper.
based on the applicant's skills. Once these training plans were devised, a party not employed by the local training office randomly assigned two thirds of the women within each training plan to the treatment group. The remaining one third were placed in the control group and prevented from receiving any JTPA service for 18 months, although they could not be prevented from receiving similar training from alternative sources. The experiment's designers randomly assigned applicants after local sites had devised training plans, so as not to disrupt the normal programme procedures and to evaluate JTPA as it actually operated in the field.

In this paper we examine how JTPA affected the employment histories of more than 2,600 adult women who programme administrators recommended to receive training services that included CT. As expected with an experimental design, the mean characteristics of the treatments and controls assigned to this training plan were nearly the same. As shown by the first two columns of Table 1, the average age of treatments and controls at the baseline was approximately 32, about one half of the sample were high school dropouts, one third were African-American, had children under 5 years old, or had been on welfare for at least 2 years prior to the baseline. Approximately one-fifth of the sample had never worked for pay. In addition, at the baseline 40 percent of participants in the experiment were receiving welfare (AFDC) and 65 percent were receiving some form of public assistance including food stamps. Finally, 84 percent were unemployed when they applied to JTPA, while 16 percent had been employed during the baseline month.

Comparisons between the treatments' and controls' employment rates show that the opportunity to receive this training improved the treatments' employment prospects. As shown by Figure 1, eighteen months after the baseline the treatments' employment rates are approximately four percentage points greater than those of the controls. This difference is statistically significant and appears to increase somewhat with time.

These treatment effects do not measure the impact of receiving CT or of receiving any other JTPA service. Instead, they measure the impact of the "intention to treat" or of the opportunity to receive JTPA–CT services. For the adult women assigned to the treatment group, a relatively large percentage received JTPA services. During the 18 month period following the baseline, 73 percent received some JTPA services and 89 percent of these participants received CT (Kemple, Doolittle, and Wallace (1993)). Among the controls, 34 percent reported receiving similar types of training during the sampling frame. Indeed, approximately 88 percent of the total hours spent in training by each experimental group was spent in classroom training. Further, both treatment and control trainees received the vast majority of this training from a two-year community college or a vocational school. Because treatment status differs from training status for a substantial number of women, we present separate sets of estimates of the treatment effects and the training effects below.

2.3. Treatments' and controls' empirical hazard rates

The National JTPA study showed that access to JTPA–CT raised the short-term employment rates of adult women. Given this result, we ask whether this rise resulted from longer spells of employment, shorter spells of unemployment, or both. Understanding how these programmes raise the economically disadvantaged's employment rates can aid policy

3. According to administrative records, the median time that the treatments spent enrolled in JTPA services is approximately 5 months; the self-reported measures (which are used in this paper) of the time spent in training are somewhat smaller.
TABLE 1
Means of treatments’ and controls’ demographic characteristics

<table>
<thead>
<tr>
<th></th>
<th>Full sample controls</th>
<th>Fresh unemployment spells controls</th>
<th>Fresh employment spells controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>A. Characteristics used in analysis:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at baseline</td>
<td>31.7</td>
<td>31.6</td>
<td>31.7</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.3)</td>
<td>(0.2)</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>11.3</td>
<td>11.3</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>High school dropout</td>
<td>0.49</td>
<td>0.50</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Kids under 4 years old</td>
<td>0.35</td>
<td>0.32</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Never married</td>
<td>0.34</td>
<td>0.39</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Married at baseline</td>
<td>0.21</td>
<td>0.21</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Black</td>
<td>0.33</td>
<td>0.33</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.16</td>
<td>0.14</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>B. Employment and welfare histories:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never worked for pay</td>
<td>0.20</td>
<td>0.21</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Weeks worked last year</td>
<td>16.64</td>
<td>16.22</td>
<td>18.46</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.73)</td>
<td>(1.09)</td>
</tr>
<tr>
<td>On AFDC at baseline</td>
<td>0.45</td>
<td>0.49</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>On AFDC &gt; 2 years</td>
<td>0.33</td>
<td>0.39</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Never on AFDC</td>
<td>0.39</td>
<td>0.37</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>On any public assistance at baseline</td>
<td>0.59</td>
<td>0.63</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Number of women in sample</td>
<td>1,584</td>
<td>747</td>
<td>686</td>
</tr>
<tr>
<td></td>
<td></td>
<td>318</td>
<td>1,145</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>502</td>
</tr>
</tbody>
</table>

Notes: See the appendix for details on how the sample was constructed. The standard errors are in parentheses.

makers and programme administrators when they modify existing programmes or when they design new programmes. One goal of training providers is to reduce the propensity of disadvantaged women to have many short spells of employment interspersed with sometimes long spells of nonemployment. Therefore, an objective of these programmes is to increase participants’ skills so that they can hold on to their subsequent jobs longer and thereby be in a position to acquire additional skills through their employer. In their study of a successful WE programme (targeted toward a similar though somewhat more impoverished population) Ham and LaLonde (1996) found that the significant rise in employment rates resulted entirely from the programme’s effects on the length of trainees’ employment spells. A second reason for examining how successful training programmes affect employment and unemployment durations is that the estimates can be used to predict the impact of training beyond the short sampling frame that is usually available to programme evaluators.
To begin our analysis, we compare the empirical hazard rates of the treatments and controls out of the following four types of spells:

(i) unemployment spells in progress at the baseline (interrupted unemployment spells);
(ii) unemployment spells that begin after the baseline (fresh unemployment spells);
(iii) employment spells in progress at the baseline (interrupted employment spells);
(iv) employment spells that begin after the baseline (fresh employment spells).

The differences between the treatments' and controls' empirical hazard rates suggest that training's impact on employment rates resulted from increasing the transition rates from unemployment. As shown by Figures 2a and 2b, the treatments usually have greater hazards rates out of unemployment. (During the first month of their interrupted spells, their hazard rates are lower than those of the controls, but this shortfall may result because access to training may initially delay reentry into the work force.) By comparison, Figures 2c and 2d suggest that access to CT had no effect on the transition rates out of employment.

3. ECONOMETRIC METHODOLOGY

3.1. Estimating the treatment effects

Because random assignment took place at the baseline when an individual was in the midst of an interrupted spell, there is no reason to believe that in the samples of fresh employment or unemployment spells an individual's treatment status is independent of her unobservable characteristics. We now develop a framework for examining whether
the foregoing comparisons of treatments' and controls' transition rates from fresh employment and unemployment spells are misleading because of dynamic selection bias resulting from differential sorting by treatments and controls into these spells. To do this, we jointly estimate the parameters of the hazard functions for the interrupted and fresh spells by maximizing the appropriate likelihood function while allowing the unobserved heterogeneity terms to be correlated across the spells. Note that we are not focusing on the bias that occurs (even with random assignment) within a spell because of unobserved heterogeneity (Ridder and Verbakel (1984)).

In forming this likelihood function, it is extremely difficult to estimate the exact contribution of the labour market state at the baseline and the time remaining in the (interrupted) spell in progress at the baseline. As a result, we adopt Heckman and Singer's
approximate solution to this problem, and specify a hazard function for the interrupted employment (unemployment) spell that differs from the fresh employment (unemployment) hazard in terms of the parameters and the heterogeneity term (Heckman and Singer (1984a), p. 127). 4

4. The hazard functions for the interrupted spells are conditional on the labor market state at the baseline. Initially we had hoped to separate out the initial state from the hazard for the interrupted spells, as suggested by Heckman and Singer ((1984a), p. 125, equation 75). To do this we need a variable that helps predict the initial state but does not enter the hazards for the interrupted spells, and unemployment rates prior to the time of entry would fulfill this role. Unfortunately, there is a break in the unemployment series and values prior to January 1988 are not comparable with those in that month or later. Since individuals enter training during the period January 1988 to late 1989, there also is a break in the series for the lagged demand variables at the baseline. Thus we estimate the interrupted spell hazards conditional on the initial state.
Specifically, we define the hazard function for leaving spell type \( j \) during month \( k \) by using a logit function

\[
\lambda_j(k | \theta_j) = \frac{1}{1 + \exp(-z_j(k))},
\]

where

\[
z_j(k) = \alpha_j + \beta_{1j}x(k) + \beta_{2j}TR + h_j(k + \tau) + \theta_j
\]

for spell types \( j = u' \) (an interrupted unemployment spell), \( e' \) (an interrupted employment spell), \( u \) (a fresh unemployment spell), and \( e \) (a fresh employment spell). To streamline notation, we do not index functions or variables by an individual subscript. The term \( \alpha_j \) is a constant for spell type \( j \) and \( \theta_j \) denotes the unobserved heterogeneity in that spell. The term \( TR \) is a dummy variable equal to one if the individual is in the treatment group which is offered JTPA training. Because we use the Heckman-Singer (1984a) approach to characterize spells in progress at the baseline (as defined by equations (1) and (2)), random assignment ensures that \( TR \) is independent of the unobserved heterogeneity in the interrupted spells.

The term \( x(k) \) represents a vector of additional explanatory variables. A listing of these variables is provided in the notes to Table 2. Some of these variables, such as monthly local unemployment rates and age, change with time and thus help to identify the parameters of the joint likelihood.\(^5\) Our use of the time-changing variables gives us an exclusion restriction in the sense that the values of these variables from the interrupted spells affect the fresh spells only through the selection process. We also included a dummy variable for the baseline month (when volunteers are assigned to the treatment and control groups), and an interaction term between this variable and a woman’s treatment status. These two additional variables had statistically significant coefficients in the interrupted spell hazards.

The term \( h_j(k + \tau) \) captures the duration dependence in the respective hazard. For interrupted and fresh spells, \( k \) is the time spent in the spell since the baseline. For an interrupted spell, \( \tau \) is the number of months spent in the spell before the baseline, while for a fresh spell \( \tau \) equals zero. (For notational ease we do not condition explicitly on \( \tau \) when writing the hazard.) To ensure that our results are not biased by using a too restrictive form of duration dependence, we use a fourth order polynomial in log duration.\(^6\) We control for the pre-baseline duration of the interrupted spells, \( \tau \), because Ham and LaLonde (1996) found that this variable was essential for identification in their multipell likelihood function with correlated heterogeneity. We also find this to be true when using a similar likelihood function for our JTPA data. Therefore, prior duration in the current labour market state can be an important identifying variable in multiple spell duration models. We believe that this data should be obtained when collecting event history data, especially because it usually can be acquired at a relatively low cost.\(^7\)

---

5. We also estimated our model (without heterogeneity) for the cases where: (i) we replaced the local unemployment rate with 15 site dummy variables and where (ii) we included both the local unemployment rate and 15 site dummy variables. These alternative specifications had no impact on the other estimated coefficients.

6. We tried a number of specifications for the duration dependence structure for the case where we do not allow for unobserved heterogeneity. Once we allowed for a reasonable degree of flexibility the results were not sensitive to changes in the specification. For a discussion of these issues see Moffitt (1985), Kennan (1985), Ham and Rea (1987), Meyer (1990), Narendranathan and Stewart (1993) and Baker and Rea (1995).

7. For researchers who do not have data on prebaseline duration, it is useful to note that we obtained identification when we did not condition on prebaseline duration by constraining the intercept terms in (2) to equal zero. Moreover, this modification did not affect our empirical findings. Ham and LaLonde (1996) also found this result.
Conditional on the unobserved heterogeneity, the probability that a spell of type $j$ lasts longer than $t-1$ months is given by the survivor function

$$S_j(t-1 | \theta_j) = \prod_{s=1}^{t-1} [1 - \lambda_j(s + \tau | \theta_j)].$$

(3)

The conditional density of a spell that lasts $t$ months is given by

$$f_j(t | \theta_j) = \lambda_j(t | \theta_j) S(t - 1 | \theta_j).$$

(4)

Because the likelihood function for this type of data is discussed in Ham and LaLonde (1996), here we simply describe its construction for two examples. First, consider the employment history depicted in Figure 3. This woman has been unemployed at the baseline for $\tau$ months and remains unemployed for $ru$ additional months before finding a job. She then experiences an employment spell lasting $te$ months, followed by an unemployment spell that is censored after $tu$ months at the end of the sample period (date $T$). (Note that she enters CT in period $tr^*$.) Her contribution to the likelihood is

$$L = \int_{\Theta} f_u(ru | \theta_u) f_e(te | \theta_e) S_u(tu | \theta_u) dG(\theta_u, \theta_e, \theta_u),$$

(5)

where $G(\theta_u, \theta_e, \theta_u)$ is the distribution function of the heterogeneity terms.

In Figure 4 we consider a woman who is employed at the baseline. She has a transition to unemployment after $re$ months, remains unemployed for $tu1$ months, is employed for
te months, and finally is unemployed during the remaining \( tu2 \) months of the sample period. Her contribution to the likelihood is

\[
L = \int_{\Theta} f_e (re | \theta_e) f_d (tu1 | \theta_u) f_e (te | \theta_e) S_u (tu2 | \theta_u) d\bar{G} (\theta_e, \theta_u, \theta_e).
\] (6)

We must specify the heterogeneity distribution to estimate the parameters of the likelihood function. It is important to emphasize that because we have random assignment among volunteers assigned to JTPA–CT, we are specifying the heterogeneity distribution conditional on volunteering and are not specifying the heterogeneity distribution for the population of low income workers. (The characteristics of volunteers would likely change if, for example, local economic conditions changed.)

In our empirical work below, we present separate estimates based on three different assumptions about the underlying heterogeneity. First, we assume that there is no unobserved heterogeneity. Second, we assume that the heterogeneity terms are independently distributed across the different hazard functions, with each being drawn from a discrete distribution. Under either of these assumptions there is no dynamic selection bias in the fresh spells caused by a differential sorting of treatment and controls into these spells. Thus the parameters for the hazard functions can be estimated separately across the different types of spells.

Our primary specification of the heterogeneity distribution for addressing dynamic selection bias is a one-factor loading model

\[
\theta_e = c_1 \theta, \quad \theta_u = c_2 \theta, \quad \theta_v = c_3 \theta, \quad \theta_e = c_4 \theta,
\] (7)

where \( \theta \) is drawn from a discrete distribution. We leave the mean and variance of \( \theta \) unrestricted and instead impose the normalizations \( c_4 = 1 \) and \( \alpha_e = 0 \) in equation (2). We let the data determine the number of support points, and thus the estimated standard errors are conditional on the estimated number of support points (Heckman and Singer (1984b)).

3.2. Estimating the training effects

We also estimate the impact of actual participation in training on the hazard rates from interrupted and fresh employment and unemployment spells (i.e. the training effects). We replace the treatment dummy, \( TR \), in equations (1) and (2) with a variable indicating whether the individual enters training in or before the current month. Specifically, \( AT(k) \) equals zero for all months before a woman enters training, and equals one in the month she enters training and in all subsequent months. The training dummy \( AT(k) \) differs from the treatment dummy \( TR \) in three ways. First, some members of the treatment group do not show up for training. For these women, \( TR = 1 \), but \( AT(k) = 0 \) for all periods. Second, \( TR \) is constant during the sampling frame but \( AT(k) \) can change over this period. Third, some members of the control group obtain training from a non-JTPA source. For these women, \( TR = 0 \) but \( AT(k) = 1 \) once training begins. We assume that the training received by the controls is comparable to JTPA–CT since this is borne out by the data, as discussed in Section 2.2.

In estimation we must take into account an institutional feature of our data. For each woman the first month of our data is the baseline month and training begins (at the earliest) during the second month. During the baseline month, \( AT(k) \) equals zero
exogenously for all members of our sample, and we can no longer interact the first month dummy with the training dummy as we do when we estimate the treatment effect.8

Unlike treatment status, actual participation in training is not determined by random assignment and the participation dummy $AT(k)$ must be treated as endogenous. We can address this endogeneity issue because the treatment dummy $TR$ is now excluded from (2), and it helps predict the training dummy $AT(k)$ (Angrist and Imbens (1991), Imbens and Angrist (1994), and Heckman, Smith and Taber (1994)).

To address this endogeneity issue, we specify the transition rate for leaving the “no training” state and entering training. We expect that this transition rate will depend on a woman’s employment status at the baseline. We estimate the model only for those unemployed at the baseline, because we do not have enough data to estimate a separate transition rate into training for those employed at the baseline. We assume that the hazard for entering training in month $m$ after the baseline is given by

$$\lambda_{et}(m|\theta_{et}) = \frac{1}{1 + \exp(-z_{et}(m))},$$

(10)

where

$$z_{et}(m) = \alpha_{et} + \beta_{et}x(m) + \gamma_{et}TR + h_{et}(m) + \theta_{et}.$$  

(11)

For the woman in Figure 3, the probability that she enters training in month $tr^*$ is

$$f_{et}(tr^*|\theta_{et}) = \lambda_{et}(tr^*|\theta_{et}) \prod_{s=2}^{r^*} [1 - \lambda_{et}(s|\theta_{et})].$$  

(12)

(Recall that training cannot begin in period 1.) Her contribution to the likelihood function is

$$L = \int_{\Theta} f_{et}(tr^*|\theta_{et}) f_u(tr|\theta_u)f_e(te|\theta_e)S_u(tu|\theta_u) dG^*(\theta_{et}, \theta_u, \theta_e, \theta_u).$$  

(13)

When estimating the training effect, we consider four specifications of the unobserved heterogeneity distribution across the hazards: (i) no heterogeneity; (ii) independent heterogeneity across all the hazards; (iii) correlated heterogeneity terms across the employment and unemployment hazards as described by equation (7), which are independent of the unobserved heterogeneity in the hazard for entering training; and (iv) correlated heterogeneity across the hazards, analogous to that described by equation (7). Specification (iii) addresses dynamic selection bias in the fresh employment and unemployment spells while (iv) allows for this bias and for the endogeneity of training participation.

4. ESTIMATES OF THE TREATMENT AND TRAINING EFFECTS

We first present estimates of the treatment effects on the transition rates from interrupted and fresh unemployment and employment spells. We then present the estimates of the training effects. In the text we present estimates only of the treatment and training effects. The other parameter estimates are contained in Appendix Tables A2 and A3.

8. We considered a specification for the interrupted unemployment hazard which included: (i) a baseline month dummy variable and (ii) its interaction with treatment status. Identification is not a problem here, because treatment status only appears during the baseline month, and the training dummy is exogenously zero during the baseline. After the baseline month, the actual training dummy becomes endogenous, but the treatment dummy is excluded from the hazard. Our estimates based on this alternate specification were essentially identical to those reported in Table 3.
We usually could identify only two points of support when we assumed independent heterogeneity distributions across the hazard functions. In the few cases in which we found a third point of support, the estimated treatment and training effects were insensitive to adding this point of support. Consequently, we report the two point of support estimates for this heterogeneity specification. We found three points of support when estimating specification (iii) for the treatment effect (i.e. the one-factor model given by equation (7)) and report these estimates in the text. When estimating the training effect, we found three points of support when a one-factor model is used for the unobservables in the unemployment and employment hazards. This occurred both for the case when the heterogeneity term for entering training is assumed independent of these unobservables (specification (iii)), and the case where it is assumed to be correlated with these unobservables (specification (iv)).

The estimated coefficients for the other variables in our models are similar among our different specifications of the treatment and training hazards. As a result, we present only the full set of estimates for the correlated heterogeneity specifications in Appendix Table A2 (treatment effects) and Appendix Table A3 (training effects). As shown by columns (1) and (2) of Tables A2 and A3, more years of schooling raised transition rates out of unemployment. By contrast, the transition rates out of unemployment were lowered by the presence of children under 4, never having been married, growing older, and higher local unemployment rates. For the employment spells, transition rates out of employment increased when job holders were high school dropouts or black, and decreased when job holders were better educated.

To see whether these characteristics also influenced the magnitude of the treatment effect, we interacted the treatment dummy with each of the explanatory variables listed in Panel A of Table A1. We did not find any statistically significant interactions. As a result, in the remainder of this paper we assume that there is no heterogeneity in the treatment effect.

4.1. The treatment effect on the employment and unemployment hazards

We first consider the treatment effect in the unemployment hazard rates. As shown by Panel A of Table 2, the results for the interrupted unemployment spells indicate that after the first month, the opportunity to participate in CT leads to increased transition rates out of unemployment. The coefficient of 0.132 in column (3) implies that women with access to CT spent on average six fewer months in their interrupted unemployment spells than did members of the control group. As seen from the estimated standard error, this effect is statistically significant at the 10 percent level.

9. The coefficients on the interactions are sizable in some cases, but in these cases the standard errors are also large. Thus we cannot identify these interaction effects from the JTPA-CT data.

10. As shown by Table A2, the baseline month effect suggests that both groups leave unemployment more quickly in the baseline month than in later months, with the controls leaving more rapidly than the treatments. These results are consistent with the individuals unemployed at the baseline becoming more active in the labor market as they volunteer for training, with the prospect of receiving training reducing this effect for the treatments. We explored whether CT reduced transition rates in the period that most of the treatments received training, but never found any effect beyond the baseline month.

11. We only see a limited range of durations, especially in the fresh spells. However, to calculate the expected durations for the interrupted and fresh spells, as well as for the steady state employment rates, we need to predict the hazard well beyond the durations we observe in our data. In making these calculations, we assume that i) in the interrupted spells the hazard is constant from 30 months and on, and ii) in the fresh spells the hazard is constant from 15 months and on.
### TABLE 2

The impact of treatment status on transition rates out of employment and unemployment spells

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Interrupted unemployment spells:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment status</td>
<td>0.134</td>
<td>0.137</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.071)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-5.184.6</td>
<td>-5.182.6</td>
<td>-12.281.7</td>
</tr>
<tr>
<td><strong>B. Fresh unemployment spells:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment status</td>
<td>0.163</td>
<td>0.178</td>
<td>0.169</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.097)</td>
<td>(0.091)</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-2.334.3</td>
<td>-2.332.0</td>
<td>-12.281.7</td>
</tr>
<tr>
<td><strong>C. Interrupted employment spells:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment status</td>
<td>-0.123</td>
<td>-0.105</td>
<td>-0.098</td>
</tr>
<tr>
<td></td>
<td>(0.240)</td>
<td>(0.273)</td>
<td>(0.284)</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-647.2</td>
<td>-646.9</td>
<td>-12.281.7</td>
</tr>
<tr>
<td><strong>D. Fresh employment spells:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment status</td>
<td>-0.050</td>
<td>-0.067</td>
<td>-0.077</td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.080)</td>
<td>(0.087)</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-4.140.8</td>
<td>-4.128.2</td>
<td>-12.281.7</td>
</tr>
</tbody>
</table>

**Specification:**
- Controls for unobserved heterogeneity: No, Yes, Yes
- Correlated among spells: No, Yes

**Notes:** Estimated coefficients are based on equations (5) and (6) in the text. Each specification also includes controls for years of schooling, high school dropout status, marital status, number of children under 4, race, unemployment rate in the locale, and a fourth order polynomial in log duration of the current spell. The estimated coefficients associated with these variables were similar among the different specifications. The estimates in column (3) are based on the correlated heterogeneity specification depicted in equation (7). Therefore, the log likelihood in column (3) contains the contribution of all four spells. A complete set of estimates for this specification is found in the Appendix. The standard errors are in parentheses.

Given our specification of the interrupted spell hazards, random assignment ensures that TR is independent of the unobserved heterogeneity in these hazard functions. As a result, consistent estimates of the interrupted unemployment (employment) hazards can be obtained from data only on interrupted unemployment (employment) spells, i.e. there is no need for joint estimation. Thus it is reassuring that the treatment effect is not affected by moving from column (2) (independent heterogeneity) to column (3) (correlated heterogeneity), since incorrectly imposing the heterogeneity structure (7) could bias the results. (This is the same reasoning as that used by Hausman (1978) in proposing a specification test comparing 2SLS and 3SLS estimates.) Because the results in columns (1) (no heterogeneity) are quite close to those in columns (2) and (3), there is no evidence of the standard bias from ignoring unobserved heterogeneity within a spell.

As shown by Panel B of Table 2, the availability of JTPA–CT services also raises the hazard rate out of the fresh unemployment spells. The coefficient in column (3) indicates that the expected durations of treatments’ fresh unemployment spells were approximately four months shorter than those of the controls. Our estimate of the treatment effect is insensitive to our specification of the unobserved heterogeneity. This result indicates that dynamic selection bias is not a problem in estimating the fresh unemployment hazard.
Our estimated training effects on unemployment durations may seem implausibly large, given the modest four percentage point gap between the two experimental groups' employment rates shown in Figure 1. However, it is worth noting that 18 months after the baseline, 33 percent of the treatments and 38 percent of the controls were still in their interrupted unemployment spells. For these women, the treatment effect in the fresh unemployment spell could not contribute to the treatments' higher employment rates (because they did not reach a fresh unemployment spell).

Our finding of shorter unemployment durations is consistent with what we would expect if training enhanced participants' job search skills. Although women in JTPA–CT were much more likely to receive classroom training than other JTPA services, many participants also received some job search training (JSA). Administrative records indicate that the percentage of enrollees receiving at least some JSA services was significant in five of the 16 sites. Because past studies indicate that JSA is often a cost effective service, we were concerned that our findings resulted from effective JSA services as opposed to CT (LaLonde (1995)). To examine this possibility, we excluded women from the five sites in which significant numbers received JSA and re-estimated the model with no heterogeneity. Indeed, the resulting estimates for the fresh unemployment spells indicated that the increased transition rates were generated entirely from sites offering JSA. But this result did not hold for the interrupted unemployment spells, as we found that the estimated treatment effect was larger in the sites that did not offer JSA to a large number of treatments.

We now turn to our analysis of the employment spells. As shown by Panel C of Table 2, the availability of CT reduced the probability of leaving the employment spell in progress at the baseline. However, the estimated standard errors are relatively large and none of these estimated effects is statistically significant, perhaps because only fourteen percent of the sample experiences an interrupted employment spell. This result is insensitive to our specification of the unobserved heterogeneity distribution.

As shown by the last panel of Table 2, the estimates of the treatment effect in the fresh employment hazards also are insensitive to the treatment of unobserved heterogeneity. This finding indicates that dynamic selection bias also is not a problem in estimating this hazard. The coefficient of −0.077 (in column 3) implies that the availability of CT reduced the probability of leaving a fresh employment spell, and thereby raised the treatments' employment durations by approximately two months. However, the estimated standard errors in the panel indicate that none of the estimated effects is statistically significant, even though we have data on a relatively large number of fresh employment spells.

Our results for the treatment effects are robust to two specification changes. First, we allowed for a separate one-factor heterogeneity distribution (analogous to equation 7) for (i) those in an interrupted unemployment spell at the baseline and (ii) those in an interrupted employment spell at the baseline. Second, we allowed both the fresh spell hazards and the unobserved heterogeneity distribution (7) to depend on the baseline employment state; in this case we could only estimate the model for those in an interrupted unemployment spell at the baseline. (Using only the data on those in an interrupted unemployment spell does not cause a selection problem because a woman's experimental status is independent of her labour market state at the baseline by random assignment.) Our results are unaffected by these modifications.

12. As shown by Table A2, there is a large effect of the baseline month on the hazard, but this effect does not differ by treatment status. This result is consistent with individuals volunteering for training as they lose their jobs.
Finally, to examine the plausibility of our parameter estimates, we construct "goodness of fit tests" based on the difference between the predicted employment rates generated by our model and the actual employment rates presented in Figure 1 (Heckman and Walker 1990). Chi-square statistics for the difference between treatments' and controls' predicted and actual employment rates at 18, 24, and 30 months were never statistically significant for any of the models.

To project the long-term "steady-state" employment rates implied by the fresh spell results, we calculate the ratio of the expected employment duration to the sum of the expected durations of the fresh employment and unemployment spells for each group. Our parameter estimates imply that the expected durations of treatments' fresh employment and unemployment spells are 28.6 and 16.1 months, respectively. For the controls, these durations are 26.2 and 20.0 months, respectively. Based on this calculation, we estimate that the treatments' long-term employment rate is 0.64 and the controls' rate is 0.57, resulting in a difference of 7 percentage points.

4.2. The training effect on employment and unemployment hazards

Because some treatments were "no shows" and some controls received similar training elsewhere, we expect that the estimated training effects will be somewhat larger than the estimated treatment effects. From Panel A of Table 3 we see that the estimates of the training effect for the interrupted unemployment spells are virtually identical across the first three heterogeneity specifications which treat training status as exogenous. The estimate in column (1) of Table 3 implies that the expected duration of an interrupted unemployment spell for a woman who participates in training is 6.4 months shorter than that for a woman who does not participate in training. This benefit from training is comparable to our estimated treatment effect reported above in Section 4.1.

When we allow for the endogeneity of training in column (4), the coefficient essentially doubles. Moreover, this coefficient is relatively precisely determined, indicating the value of having treatment status as an explanatory variable in the training entry hazard. For policy purposes, this increase in the training effect is substantial. The estimate in column (4) implies that a woman who participates in training has an expected duration that is 16.9 months shorter than that of an individual who does not participate in training. Thus, there appears to be adverse sorting into training, and that ignoring this sorting causes one to underestimate the effect of training. Moreover, once we allow for the endogeneity of training status, the estimated training effect is indeed larger than the corresponding treatment effect.

As shown by Panel B of Table 3, allowing for the endogeneity of training has no impact on our estimates of training effect in fresh unemployment spells. These estimates imply that for someone who participates in training, the expected duration of a subsequent fresh unemployment spell is eight months shorter than that for someone who does not participate in training. Further, this training effect is about twice as large as the treatment effect of 4 months.

Allowing for the endogeneity of training participation influences the estimate of the training effect in the fresh employment spells. As shown by Panel C of Table 3, the training effect in the fresh employment spells increases by a factor of 1.5 for the treatments and 1.3 for the controls.

13. We should also note that as shown by Table A3, the estimates of the heterogeneity distribution suggest that individuals whose unobservable characteristics make them more likely to enter training also make them less likely to leave an interrupted unemployment spell and more likely to leave employment. Of course, the estimates of the heterogeneity distribution should be viewed with some caution given Heckman and Singer's (1984b) results.
null

null

null

null
5. WHEN IS DYNAMIC SELECTION INTO FRESH SPELLS IMPORTANT IN EVALUATING TRAINING PROGRAMS?

We find no evidence in the JTPA–CT data that differential sorting by the treatments and controls into fresh spells biases the estimation of these hazards. This finding contrasts with Ham and LaLonde's (1996) study of the National Supported Work–Work Experience (NSW–WE) programme. To investigate why dynamic selection bias is important in NSW–WE but not JTPA–CT, we compare the characteristics of a stylized WE programme and a stylized CT programme.

The programmes share the following characteristics. Because the programmes serve the economically disadvantaged, training volunteers are usually unemployed at the baseline and many have been unemployed for a substantial period of time. For simplicity, we assume that everyone is in an interrupted unemployment spell at the baseline. Second, we assume that we follow treatments and controls until i) they experience both a fresh employment spell and a fresh unemployment spells or ii) the sampling frame ends. We also assume that the sampling frame in each case is two years. Since this sampling frame is relatively short, we would not expect to observe both a fresh employment and unemployment spell for every woman in the sample. Third, we assume a stationary environment.14

The difference between the programmes can be characterized as follows: Treatments assigned to the WE programme leave unemployment and enter a subsidized job (often in a sheltered environment) for one year. All treatments return to unemployment after training ends and are followed for one year further.15 By contrast, the controls continue in their interrupted unemployment spells and are followed for a total of two years. Unlike participants in the WE programme, treatments assigned to the CT programme participate in training on a part-time basis and can search for (and hold) a regular job while acquiring training; participation in JTPA–CT does not automatically end an individual's interrupted unemployment spell. All individuals in CT are followed for two years.

When evaluating the WE programme, dynamic selection bias may arise in the following manner. The programme ends all of the treatments’ interrupted unemployment spells and then places them in a fresh unemployment spell when they leave training. Thus, the selection process dictates that the treatments observed in fresh unemployment spells are the stock of treatments who were unemployed at the baseline. For a control to reach a fresh unemployment spell, she must first exit from the unemployment spell in progress at the baseline and then complete an employment spell. As a result, we will observe only a fraction of the controls in a fresh unemployment spell during a short sampling frame. The

14. Relaxing this assumption would complicate our discussion, because now observing a fresh spell for every treatment and control is not sufficient to avoid dynamic selection. Time-changing explanatory variables affect when individuals experience a fresh spell and can induce a correlation between these variables, the unobserved heterogeneity and the other explanatory variables (including treatment status) in the fresh spells.

15. In practice, some treatments will not experience a fresh unemployment spell during the sample frame because they do not show up for training and remain unemployed, or because they leave training and enter a fresh employment spell. Because 81 percent of the treatments in Ham and LaLonde's (1996) data experienced a fresh unemployment spell, we believe our assumption captures the essential features of the NSW WE program while simplifying the discussion.
sample of controls who experience a fresh unemployment spell constitutes a flow sample, first out of the interrupted unemployment spell and then out of the fresh employment spell.

In general, we would expect the flow and stock samples to have different distributions of unobserved characteristics (Cheshier and Lancaster (1983)). In a WE programme, we expect controls who leave interrupted spells on their own to have above average ability. As a result, we compare the stock sample of treatments with above average controls. Consistent with this conjecture, Ham and LaLonde report that the NSW–WE controls who experienced a fresh unemployment spell had nearly twice as much prebaseline employment experience as did the treatments. Further, they found substantial dynamic selection bias in their estimates of the treatment effect in fresh unemployment spells.

There also is the potential for dynamic selection bias to affect the evaluation of fresh employment spells that follow the WE programme. However, unlike the fresh unemployment spells, it is not clear a priori whether it is harder for a treatment or a control to reach a fresh employment spell by the end of the sampling frame. Assume that WE does not affect unemployment duration, as found by Ham and LaLonde for NSW–WE. However, even with this assumption, the negative duration dependence in the interrupted and fresh unemployment spells (as found in the NSW–WE data) indicates that a control is less likely to leave her interrupted unemployment spell during any given month than a treatment is to leave her fresh unemployment spell. This occurs because the controls are at much higher duration levels in the interrupted spells. This effect suggests that we also should see higher ability controls in the fresh employment spells. (If training had raised the hazard rate out of unemployment, this would have further skewed the composition of treatments and controls with fresh employment spells.) On the other hand, the controls have two years in which to leave their interrupted unemployment spells, while the treatments only have one year to enter a fresh employment spell (since they are in training for one year). This second effect may counter-balance the first effect. In fact, Ham and LaLonde found that in NSW–WE a similar number of treatments and controls experienced fresh employment spells, although these controls had two-thirds more prebaseline employment experience than the treatments. However, Ham and LaLonde find no evidence of dynamic selection bias in their analysis of fresh employment spells for NSW–WE.

In contrast to the WE programme, participation in the CT programme does not cause the treatments’ interrupted unemployment spells to end. Therefore, unlike the case of the WE programme, the samples of treatments and controls entering fresh employment spells are both flow samples from interrupted unemployment spells. Whether there is dynamic selection bias in evaluating the effect of CT in fresh employment spells depends upon whether treatment status affects the exit rate out of the interrupted unemployment spells. If training raises these exit rates, we would expect to observe more treatments with a fresh employment spell, and we also would expect that treatments will be (on average) less able than the controls. A sufficient condition for this negative correlation between treatment status and unobserved heterogeneity in the fresh employment hazard is that the heterogeneity term be multiplicatively separable in the interrupted unemployment hazard, such as in the case of a proportional hazard. This raises the possibility that dynamic selection bias arising from the sorting on unobservables into the fresh employment spells will lead us to underestimate the treatment effect on the corresponding hazard.

16. The bias may be in the opposite direction for other specifications of the hazard. We demonstrate these results in Appendix Two. Note that our logit specification in discrete time approaches the proportional hazard function as the time period grows small (Lancaster and Nickell (1980)).
Indeed, in our JTPA–CT sample a somewhat larger fraction of treatments experience fresh employment spells (72 percent) than is the case for the controls (67 percent). However, as shown by columns (5) and (6) of Table 1, the mean characteristics of treatments and controls with such spells are nearly the same. This result holds even for "endogenous" pre-baseline characteristics such as time on welfare, weeks worked during the previous year, or whether the woman had ever held a regular job. Our results indicate that the differences between the flow samples of the treatments and controls into fresh employment spells are too subtle to affect the estimate of the treatment effect on fresh employment spells in JTPA–CT.

There is also the potential for dynamic selection bias in estimating the effect of CT in the fresh unemployment spells. Because more treatments entered a fresh employment spell, and assuming that CT has no effect on employment duration (as we find in the JTPA–CT), the above argument suggests that there may also be a negative correlation between treatment status and the unobserved heterogeneity in the sample entering a fresh unemployment spell. However, in our JTPA–CT sample we find that 43 percent of both the treatments and controls experience a fresh unemployment spell. Moreover, as shown by columns (3) and (4) of Table 1, these treatments' and controls' endogenous characteristics do not differ significantly.

Although we must be cautious about generalizing from two training programmes, we make the following conjecture based on the above comparison of the WE and CT programmes: we expect that dynamic selection is likely to be much more important for programmes that take participants out of the labour market for prolonged periods of time. In this case, the treatments subsequently leave training for a fresh spell, whereas during the sampling period only a subset of controls are able to exit their ongoing interrupted unemployment spell to experience a fresh spell. This selection process implies that the composition of treatments and controls with fresh spells is likely to differ substantially. An example of a U.S. employment programme with these characteristics is Job Corps. This programme places eligible disadvantaged youths in residential centres for up to one year and provides them with a wide range of educational, health, and training services. Accordingly, we expect that an analysis of the effect of the Job Corps on subsequent transition rates from employment and unemployment would need to account for dynamic selection bias.

At the opposite extreme, we expect this consideration to be less important when the treatment is like a "shot in the arm". In this case, the treatments' interrupted unemployment spells do not automatically end when they enter the programme. As a result, the difference between the selection process into fresh spells for the two experimental groups is likely to be much less dramatic. An example of a U.S. employment programme with these characteristics is Job Search Assistance (JSA). This programme provides only a few weeks of (part-time) job search training and career counseling.

7. CONCLUSION

We have addressed two questions in this paper. First, what is the impact of providing access to JTPA–CT on the duration of unemployment and employment spells? Second, what is the effect of actually participating in CT on the length of unemployment and employment spells? We find that the effect of being a member of the experimental treatment group raises employment rates by shortening unemployment spells, but has no effect on the length of employment spells. Further, we find no evidence of dynamic selection bias
into fresh spells in our estimates of the treatment effect in JTPA–CT. These results are quite different from Ham and LaLonde's (1996) results for NSW–WE.

We also find that actually participating in training has a larger positive effect on the exit rate from unemployment than the effect of simply being a member of the treatment group. Moreover, we find that ignoring the endogeneity of actual training status substantially underestimates its effect on labour market histories, and that we can obtain relatively precise estimates of the training effects by using experimental status to predict training status.

We argue that selection bias is a much more serious problem in the NSW–WE than JTPA–CT because the NSW–WE constitutes a much more dramatic intervention in labour market histories. Our results suggest it is probably much more important to worry about dynamic selection bias in a programme like Job Corps than in a programme such as Job Search Assistance.

APPENDIX

1. Data

1.1. Design of the experiment

Initially researchers tried to select a representative sample of the nearly 600 sites (SDAs) nationwide. However, only four of the randomly selected SDAs or their alternates agreed to participate. As a result, the study's designers selected a diverse group of SDAs which were willing to participate in a study that used random assignment. (For a discussion of the problems associated with implementing the NJS see Doolittle and Traeger (1990).) Although the NJS examined other groups, adult women recommended for CT are the focus of this paper.

Although those assigned to the control group were prevented from receiving JTPA sponsored training, they could receive other training services available in their communities. Members of the control group sometimes received a few hours of job search assistance as part of the JTPA screening process. In addition, they were given a list of non-JTPA programmes operating in their communities, but they received no referrals. This list of other training alternatives included the local Employment Service, local community colleges, technical institutes, and community agencies that provide social services.

1.2. Description of our sample

Our sample is taken from participants' Background Information Form (BIF) and from the 18 and 30 month follow-up surveys. The BIF provides information about participants' characteristics at the baseline such as schooling, work and social welfare histories, children in the household, marital status, site, experimental status, and recommended service strategy. The follow-up surveys provided information on participants' self-reported monthly employment and training status, age, race, and month of assignment. Our local labour market data is from the U.S. Department of Labour (DOL) and from various monthly issues of U.S. DOL Employment and Earnings.

Our sample includes all adult women recommended to receive CT. We excluded from our sample women whose employment status was missing on the BIF or who did not have a continuous monthly employment history covering at least the first 18 months after the baseline. In cases in which the information about employment status at the baseline was not the same on the BIF as on the follow-up surveys, we used the information from the BIF. By design, the second of the two follow-up surveys interviewed only 25 percent of the original sample, and thus only approximately 40 percent of the sample had continuous employment histories as long as 24 months and only 20 percent as long as 30 months. Our research sample included 2,671 women. When we exclude women with missing information about the duration of their interrupted unemployment and employment spells at the baseline, the sample size fell to 2,331 women. This is the sample we use for our analysis of the treatment effects in section 4.1. When we exclude women who were in an interrupted employment spell at the baseline, our sample size fell to 1,944 women. This is the sample we use for our analysis of the training effects in section 4.2.

17. An expanded version of this data appendix is available from the authors upon request.
18. The expanded data appendix (available from the authors) contains a comparison of our sample and the Bloom et al (1993) sample.
1.3. Explanatory variables

The following are the definitions of the explanatory variables used in the hazard equations:

**Treatment status**: equals 1 if assigned to the treatment group, and equal to 0 if assigned to the control group.

**Schooling**: Highest grade of schooling completed.

**No degree**: equals 1 if no high-school degree, equals 0 otherwise.

**Kids under 4**: equals 1 if children under 4 years old live with the woman at the baseline, equals 0 otherwise.

**Never married**: equals 1 if the woman had never been married at the baseline, equals 0 otherwise.

**Married**: equals 1 if the woman is married with spouse present at the baseline, equals 0 otherwise.

**Black**: equals 1 if the woman is black/nonhispanic, equals 0 otherwise.

**Hispanic**: equals 1 if the woman is hispanic, equals 0 otherwise.

**Age**: A woman's age in years.

**Unemployment**: equals the unemployment rate in the SDA for the month of the observation.

**First month**: equals 1 for the first month after random assignment, equals 0 otherwise.

1.4. Imputation of missing data

We did not discard observations with missing values for the variables Schooling, No degree, Kids under 4, Never married, and Married. Instead, we ran separate least squares regressions, for each SDA, of these variables on a woman's age and race dummies. We then replaced the missing values with the predicted values.¹⁹

1.5. Experimental impacts

The monthly experimental impacts on employment rates are reported in Table A1. The standard errors associated with the fraction employed and the difference between the fraction of treatments and controls who are employed during a given month rises after month 18 because by design the sample size becomes smaller.

---

¹⁹ The numbers of person-month observations imputed are available in the expanded data appendix available from the authors.
Parameter estimates for impact of exogenous treatment status on unemployment and employment hazards

[Full set of estimates for column (3) of Table 2]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unemployment spells (Interrupted (1), Fresh (2))</th>
<th>Employment spells (Interrupted (3), Fresh (4))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment status</td>
<td>0.132, 0.169</td>
<td>0.098, 0.077</td>
</tr>
<tr>
<td>First month</td>
<td>0.070, 0.091</td>
<td>2.952</td>
</tr>
<tr>
<td>First month* treatment status</td>
<td>0.135</td>
<td>0.314</td>
</tr>
<tr>
<td>Schooling</td>
<td>0.047, 0.069</td>
<td>0.049, 0.103</td>
</tr>
<tr>
<td>No high school degree</td>
<td>0.074, 0.018</td>
<td>0.198, 0.202</td>
</tr>
<tr>
<td>Kids under 4</td>
<td>0.024, 0.029</td>
<td>0.081, 0.032</td>
</tr>
<tr>
<td>Never married</td>
<td>0.075, 0.101</td>
<td>0.258, 0.097</td>
</tr>
<tr>
<td>Married</td>
<td>0.003, 0.111</td>
<td>0.095, 0.007</td>
</tr>
<tr>
<td>Black</td>
<td>0.065, 0.105</td>
<td>0.028, 0.197</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.039, 0.111</td>
<td>0.327, 0.045</td>
</tr>
<tr>
<td>Age</td>
<td>0.073, 0.111</td>
<td>0.325, 0.129</td>
</tr>
<tr>
<td>Age squared/100</td>
<td>0.085, 0.078</td>
<td>0.021, 0.002</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>0.093, 0.126</td>
<td>0.325, 0.129</td>
</tr>
<tr>
<td>Log duration</td>
<td>0.016, 0.025</td>
<td>0.051, 0.022</td>
</tr>
<tr>
<td>Log duration^2</td>
<td>0.468, 0.623</td>
<td>0.096, 0.024</td>
</tr>
<tr>
<td>Log duration^3/10</td>
<td>0.642, 0.271</td>
<td>0.325, 0.055</td>
</tr>
<tr>
<td>Log duration^4/10</td>
<td>0.510, 0.807</td>
<td>0.325, 0.055</td>
</tr>
<tr>
<td>Heterogeneity terms:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_1)</td>
<td>-0.932, 0.905</td>
<td></td>
</tr>
<tr>
<td>(\theta_2)</td>
<td>-2.757, 0.872</td>
<td></td>
</tr>
<tr>
<td>(\theta_3)</td>
<td>-4.185, 4.375</td>
<td></td>
</tr>
<tr>
<td>(\gamma_1)</td>
<td>-2.177, 0.784</td>
<td></td>
</tr>
<tr>
<td>(\gamma_2)</td>
<td>3.412, 0.697</td>
<td></td>
</tr>
<tr>
<td>(\alpha_i)</td>
<td>-1.125, 1.500</td>
<td>-5.465, 0.0</td>
</tr>
<tr>
<td>(\alpha_j)</td>
<td>-0.330, 0.221</td>
<td>-1.252, 1.0</td>
</tr>
<tr>
<td>(\beta_j)</td>
<td>-0.020, 0.083</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-12.281.7</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The parameter estimates are based on the specification depicted in equations (5) and (6) in the text. The specification of the unobserved heterogeneity is given in equation (7). The probabilities for the points of support are given by \(p_i = \exp(\gamma_i)/(1 + \exp(\gamma_1) + \exp(\gamma_2))\). The standard errors are in parentheses.
### Table A3

Parameter estimates for impact of endogenous training status on unemployment and employment hazards  
[Full set of estimates for column (4) of Table 3]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unemployment spells</th>
<th>Employment</th>
<th>Training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interrupted (1)</td>
<td>Fresh (2)</td>
<td>Fresh spells (3)</td>
</tr>
<tr>
<td>Training status</td>
<td>0.622</td>
<td>0.271</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.122)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>First month</td>
<td>0.182</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.189)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Schooling</td>
<td>0.034</td>
<td>0.043</td>
<td>—0.069</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.038)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>No degree</td>
<td>0.109</td>
<td>-0.074</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>(0.092)</td>
<td>(0.132)</td>
<td>(0.089)</td>
</tr>
<tr>
<td>Kids Under 4</td>
<td>-0.192</td>
<td>-0.179</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.123)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>Never married</td>
<td>-0.124</td>
<td>-0.151</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>(0.092)</td>
<td>(0.141)</td>
<td>(0.087)</td>
</tr>
<tr>
<td>Married</td>
<td>-0.051</td>
<td>0.085</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td>(0.135)</td>
<td>(0.092)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.035</td>
<td>-0.107</td>
<td>0.171</td>
</tr>
<tr>
<td></td>
<td>(0.085)</td>
<td>(0.130)</td>
<td>(0.081)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.198</td>
<td>-0.052</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.167)</td>
<td>(0.106)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.098</td>
<td>-0.059</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.032)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Age squared/100</td>
<td>0.120</td>
<td>0.058</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.069)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.069</td>
<td>-0.075</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.029)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Log duration</td>
<td>-0.134</td>
<td>0.325</td>
<td>1.954</td>
</tr>
<tr>
<td></td>
<td>(0.579)</td>
<td>(0.810)</td>
<td>(1.290)</td>
</tr>
<tr>
<td>Log duration(^2)</td>
<td>-0.144</td>
<td>-0.732</td>
<td>-1.573</td>
</tr>
<tr>
<td></td>
<td>(0.332)</td>
<td>(1.374)</td>
<td>(0.765)</td>
</tr>
<tr>
<td>Log duration(^3)</td>
<td>0.020</td>
<td>0.288</td>
<td>0.445</td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td>(0.754)</td>
<td>(0.366)</td>
</tr>
<tr>
<td>Log Duration(^4)</td>
<td>0.000</td>
<td>-0.045</td>
<td>-0.046</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.130)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Treatment status</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>First three months after baseline</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>First three months* treatment</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

**Heterogeneity terms:**

\( \theta_1 \) = -6.202  
\( \theta_2 \) = -3.538  
\( \theta_3 \) = -2.803  
\( \gamma_1 \) = -1.557  
\( \gamma_2 \) = -0.520  
\( \alpha_j \) = -3.485  
\( \gamma_j \) = -1.142  
\( \epsilon_j \) = -1.142  
\( \log \text{likelihood} \) = -13,737-9

**Notes:** The parameter estimates are based on the specification depicted in equation (13) in the text. The heterogeneity specification is given in equation (7). The estimates are based on a sample that excludes treatments and controls with interrupted employment spells. The probabilities for the points of support are given by \( p = \exp(\gamma_1)/(1 + \exp(\gamma_1) + \exp(\gamma_2)) \). The standard errors are in parentheses.
2. **Dynamic sorting when the hazard function is multiplicatively separable in the unobserved heterogeneity**

In this appendix we show that the assumption that the hazard function is multiplicatively separable in the unobserved heterogeneity is a sufficient condition to increase the proportion of “bad types” among treatments relative to controls in fresh employment spells when the sampling frame is finite and the treatment increases the hazard out of interrupted unemployment spells. By “bad types” we mean people whose unobserved heterogeneity makes it less likely that they will leave the interrupted unemployment spell, other things equal.

Let $\theta$ be the (real-valued) unobserved heterogeneity term. We assume $\theta$ is a continuous random variable with support $\Theta$. Let $g(\cdot)$ be the density function for the distribution of $\theta$. Let $d=1$ if in the treatment group and $d=0$ if in the control group. Also, let $x(s)$ denote a vector of observed characteristics and $X(t)$ the history of observed characteristics until time $t$. That is

$$X(t) = \{x(s) | 0 \leq s \leq t\},$$

where we index time so that time zero is the baseline (beginning of the sample). Also, let $\tau$ be the duration in the spell in progress at the baseline. We make the following assumptions.

- **(A1)** All individuals are in an unemployment spell at the beginning of the sample.
- **(A2)** Durations are measured in continuous time.
- **(A3)** $d$ is independent of $\theta$ and all $X(t)$.
- **(A4)** The interrupted unemployment hazard is multiplicatively separable in the heterogeneity and equals

$$\lambda(t|d,t+\tau,x(t),\theta) = \mu(\theta)\kappa(d,x(t),t+\tau)$$

$$\mu(\theta) > 0, \kappa(d,x(t),t+\tau) > 0 \forall (\theta, d, x(t), t+\tau).$$

This hazard function contains the proportional hazard as a special case (e.g. Heckman and Singer (1984a), p. 69). Specifically, we do not require that the log of the conditional hazard at $t$ is linear in functions of $t+\tau$ (duration in the spell), $(d, x)$, and $\theta$. Instead, we require only that it be linear in some function of $(t+\tau, d, x)$ and $\theta$.

- **(A5)** All individuals are followed until a fresh employment spell is completed or time $T>0$, whichever comes first.

- **(A6)** The probability of exiting an interrupted unemployment spell is at least as large when $d=1$ than when $d=0$ for any time paths of observable characteristics and all values of $\theta$ in $\Theta$ and strictly greater for some $(\theta, X(T), \tau)$.

Let $S(t|d,\tau,X(t),\theta)$ be the probability an interrupted unemployment spell has duration after the baseline greater than or equal to $t$ given $(d, \tau, X(t), \theta)$. Assumption (A4) implies that this conditional survivor for interrupted unemployment spells can be written as follows

$$S(t|d,\tau,X(t),\theta) = \exp\left[-\mu(\theta) \int_0^t \kappa(d,x(s),s+\tau)ds\right].$$

W.L.O.G. we assume $\mu(\theta) = \theta$ (note this transformation implies that the support of $\theta$ is a subset of the non-negative real numbers) so that

$$S(t|d,\tau,X(t),\theta) = \exp\left[-\theta \int_0^t \kappa(d,x(s),s+\tau)ds\right].$$

Let $f(\theta|t<T,d,\tau,X(T))$ be the conditional density of $\theta$ given $(d, \tau, X(T))$ and $t<T$, where $t$ is the duration after the baseline in the interrupted unemployment spell. Note that, given (A5), for fixed $(d, \tau, X(T))$ this is the density of $\theta$ given that a fresh employment spell is observed. By definition

$$f(\theta|\tau<T,d,\tau,X(T)) = \frac{[1-S(T|d,\tau,X(T),\theta)]g(\theta)}{\int_0^{\theta} [1-S(T|d,\tau,X(T),\theta)]g(\theta)d\theta}.$$ 

The denominator is the marginal (over $\theta$) probability that a fresh spell is observed. The numerator is the conditional (on $\theta$) probability a fresh spell is observed times the unconditional density of $\theta$.

20. We would like to thank Hidehiko Ichimura for many helpful comments and suggestions. We are responsible for any errors.

21. This assumption is not needed, but streamlines the proof; the same arguments go through for discrete or mixed distributions.
Let $\theta_1$ and $\theta_2$ be two members of $\Theta$ and assume $\theta_1 < \theta_2$. Note that $\theta_1$ is a "bad type" relative to $\theta_2$ in the sense that the unemployment survivor is greater for this type. Consider the following ratio

$$R(\theta_1, \theta_2) = \frac{f(\theta_1 | t < T, d, \tau, X(T))}{f(\theta_2 | t < T, d, \tau, X(T))}.$$  

**Proposition.** Given (A1) to (A6), $R(\theta_1, \theta_2)$ is at least as large when $d=1$ than when $d=0$ for any $\theta_1, \theta_2$ in $\Theta$ such that $\theta_1 < \theta_2$ and strictly greater for some histories $X(T)$.

**Proof.** Define the function

$$Q(d, \tau, X(T)) = \int_0^T \kappa(d(s), s + \tau) ds.$$

Note that from (A6), $Q(1, \tau, X(T))$ is greater than or equal to $Q(0, \tau, X(T))$ for any given history $X(T)$ and prebaseline duration $\tau$, with strict inequality for some $(X(T), \tau)$. Noting that $R(\cdot, \cdot) > 0$, the natural log of $R$ is

$$L = \ln(R) = \ln[1 - \exp(-\theta_1 Q(d, \tau, X(T)))] - \ln[1 - \exp(-\theta_2 Q(d, \tau, X(T))] + \ln[g(\theta_1)/g(\theta_2)].$$

Since this is a monotonic transformation, $R$ is at least as large when $d=1$ iff $L$ is at least as large when $d=1$. Note that $L$ depends on $d$ only through the function $Q$ and, by (A6) $Q$ is at least as large when $d=1$. Thus, it suffices to show that $L$ is increasing in $Q(d, \tau, X(T))$. Set $Q(d, \tau, X(T))=q$ and take the derivative of $L$ with respect to $q$ to obtain

$$\frac{\partial L}{\partial q} = \frac{\partial}{\partial q} \frac{\theta_1 \exp(-\theta_1 q) - \theta_2 \exp(-\theta_2 q)}{1 - \exp(-\theta_1 q) - \exp(-\theta_2 q)}$$

$$= -\frac{\theta_1}{\exp(\theta_1 q) - 1} - \frac{\theta_2}{\exp(\theta_2 q) - 1}.$$

To complete the proof, we need to show that the above derivative is positive. Since $\theta_1 < \theta_2$, it suffices to show that the function

$$J(\theta) = \frac{\theta}{\exp(\theta q) - 1},$$

is decreasing in $\theta$. Now

$$J'(\theta) = \frac{\exp(\theta q) - 1 - \theta q \exp(\theta q)}{[\exp(\theta q) - 1]^2}.$$

The denominator is positive. The numerator is negative since $z = \exp(y) - 1 - (y \exp(y))$ is zero when $y=0$ and strictly decreasing in $y$ for all $y > 0$. 

To show the above property does not hold for all possible hazards, we next provide a counter-example. Assume $d$ is the only observed heterogeneity and let the hazard be of the form

$$\lambda(d, \theta) = 0.001 + 0.00001\theta + 0.01d\theta^2 + 0.001d$$

Suppose $\theta_1 = 1$, $\theta_2 = 2$, $g(\theta_1) = g(\theta_2)$, and $T=1$. Then $R$ is approximately 0.29 when $d=1$ and approximately 0.99 when $d=0$. In this case (due to the term involving theta squared) the treatment helps "good types" more than it helps "bad types".

The next lemma shows that the proposition is sufficient to ensure that the expected value of the heterogeneity term is greater in the fresh spells among controls than among treatments.

**Lemma.** Assume $f$ and $g$ are real-valued density functions with common support $X$ (i.e. $f(x) > 0$, $g(x) > 0$ iff $x$ is in $X$) and finite first moments. Assume that, for any $x_1, x_2$ in $X$ s.t. $x_1 < x_2$:

$$\frac{f(x_1)}{f(x_2)} \geq \frac{g(x_1)}{g(x_2)}$$

Then

$$\int_{-\infty}^{\infty} xf(x) dx \leq \int_{-\infty}^{\infty} xg(x) dx$$

with strict equality if and only if $f$ equals $g$ almost everywhere.
Proof. By assumption
\[ \frac{f(x_1)g(x_2) - f(x_2)g(x_1)}{(x_2 - x_1)} \geq 0 \quad \forall (x_1, x_2) \in \mathbb{R}^2. \]
Integrating this over \( x_1 \) and \( x_2 \) yields the result. The above expression will be positive unless \( f(x_1)/f(x_2) = g(x_1)/g(x_2) \) almost everywhere in \( X \) when \( x_1 < x_2 \). Integrating over \( x_1 < x_2 \) yields
\[ F(x_2)/f(x_2) = G(x_2)/g(x_2) \]
where \( F \) and \( G \) are the CDFs corresponding to \( f \) and \( g \). But then
\[ \frac{d \ln (F(x))}{dx} = \frac{d \ln (G(x))}{dx}. \]
This implies \( F = G \) since they are CDFs, so \( f = g \) almost everywhere in this case. 

We are grateful to Manuel Arellano, Gerard van den Berg, Michael Cragg, Donald Deere, Hidehiko Ichimura, Guido Imbens, James Heckman, Padma Rao Sahib, Geert Ridder, Jeffrey Smith, John Xu Zheng, and three anonymous referees for very helpful discussions and comments. Participants at the Econometrics of Training Conference (Austin), the RES/CEMFI Conference (Madrid) and the Canadian Labour Economics Conference (Montreal) made very helpful comments, as did seminar participants at the University of British Columbia, University of Chicago, UCLA/RAND, Harvard/MIT, Michigan State University, Rice University, Simon Fraser University, University of Southern California, Texas A&M University, and the University of Western Ontario. The Graduate School of Business at the University of Chicago and the NSF through grant SES-9213310 generously supported this work. Stepan Jurajda provided outstanding research assistance. We emphasize that we alone are responsible for any errors.

REFERENCES


