

IZA DP No. 716

## Is Labour Market Training a Curse for the Unemployed? Evidence from a Social Experiment

Michael Rosholm  
Lars Skipper

February 2003

# Is Labour Market Training a Curse for the Unemployed? Evidence from a Social Experiment

**Michael Rosholm**

*University of Aarhus, AKF,  
SFI, CIM and IZA Bonn*

**Lars Skipper**

*University of Aarhus*

Discussion Paper No. 716  
February 2003

IZA

P.O. Box 7240  
D-53072 Bonn  
Germany

Tel.: +49-228-3894-0  
Fax: +49-228-3894-210  
Email: [iza@iza.org](mailto:iza@iza.org)

This Discussion Paper is issued within the framework of IZA's research area *Evaluation of Labor Market Policies and Projects*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor, (2) internationalization of labor markets, (3) welfare state and labor market, (4) labor markets in transition countries, (5) the future of labor, (6) evaluation of labor market policies and projects and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website ([www.iza.org](http://www.iza.org)) or directly from the author.

## ABSTRACT

### **Is Labour Market Training a Curse for the Unemployed? Evidence from a Social Experiment\***

In this paper, we investigate the impact of classroom training programmes on individual unemployment rates in Denmark. In 1994 a social experiment was conducted, where unemployed applicants for labour market training were randomised into treatment and control groups. We formulate and estimate experimental impact estimators of the effect of treatment on the treated. The experimental data is polluted by the presence of no-shows and cross-overs, which implies that traditional experimental estimators are biased. Therefore we formulate and estimate an endogenous variables model (using the randomisation indicator as a perfect exclusion restriction) and implement various matching estimators. We find – surprisingly – that classroom training significantly increases individual unemployment rates. We discuss some possible reasons for this surprising finding and some related policy issues.

JEL Classification: J64, J68

Keywords: classroom training, programme evaluation, social experiment, experimental and nonexperimental estimators

Corresponding author:

Michael Rosholm  
Department of Economics  
University of Aarhus  
Building 326  
8000 Aarhus C  
Denmark  
Tel.: +45 89 42 15 59  
Fax: +45 86 13 63 34  
Email: [mrosholm@econ.au.dk](mailto:mrosholm@econ.au.dk)

---

\* We are grateful to Jørgen Søndergaard at SFI for financial support and for kindly providing us with the data. We are grateful to colleagues at SFI, CIM and at the IZA workshop “Improving Labor Market Policies by the Evaluation of their Impact”, June 1, 2001, for valuable comments where a previous version of this paper was presented. We are particularly grateful to Astrid Grasdahl for many valuable comments.

# 1 Introduction

This paper uses data from a social experiment to evaluate the impact of classroom training on the subsequent individual rate of unemployment of those receiving training. In doing so, we exploit the information from one among only a handful of social experiments conducted in Europe on labour market issues.

Social experiments are seen by many as *the* way to estimate consistently the effect of a policy on an outcome variable of interest. The trick is to randomise the selection process itself, such that the outcomes in the control group can be used as an estimate of the unobserved counterfactual, namely, the outcome of trained individuals in the absence of training. However, as we show in this paper, even experimental data are often flawed. In the present study, the data is contaminated by the presence of no-shows (members of the treatment group who do not show up at the course they signed up for) and cross-overs (members of the control group who end up receiving the treatment they were randomised out of). The presence of no-shows, cross-overs, and also drop-outs is a well known phenomenon in social experiments, see Heckman, Hohmann & Smith (2000). However, in actual applications, their presence is rarely controlled for, in spite of the fact that neglect of them may seriously bias the estimates of the parameter of interest, simply because they constitute the very phenomenon that the experiment is designed to eliminate, namely, non-random selection.

Our goals in this paper are twofold. First, we contribute to the relatively sparse literature on experimental evaluations of labour market training programmes in Europe. In doing so, we exploit a unique data set from a social experiment conducted in Denmark in 1994. The data consists of a survey combined with information from administrative registers, which allows us to obtain very detailed labour market histories for the participants in the experiment. Second, we contribute to the literature on (experimental) evaluations by comparing several experimental estimators obtained by ignoring the process of selection into training with several non-experimental estimators, which take explicit account of the selection process. We apply a variety of identifying assumptions and corresponding experimental and non-experimental estimators. We move from the basic experimental estimator to other, simple, estimators, in order to check whether these estimators are sufficient to correct for the presence of no-shows and cross-overs. We then look at more advanced non-experimental estimators, namely, endogenous variables and matching estimators. In the case of the endogenous variables estimator, we exploit the presence of the perfect exclusion restriction, the randomisation indicator.

We find that training increases the subsequent individual rate of unemployment, relative to the counterfactual of no training, in the first months after training. This effect tends to disappear over time, though, such that six months after training, the difference between the trained and the non-trained is no longer statistically significant. To see the plausibility of this result, decompose the effect of training in two; first, there is the locking-in effect, which reduces search intensity while in training, and secondly, there is the post-programme effect, which affects the arrival rate of job offers after the completion of training. Normally, we think of the locking-in effect as increasing individual unemployment duration (hence the name), while the post-programme effect supposedly reduces unemployment duration. In the case of training which provides the trainee with formal qualifications, even the post-programme effect may become negative due to narrower job search, and/or higher reservation wages. If we think of an individual acquiring a bus-drivers' license, for example, this may induce that individual to look only for jobs as a bus driver, which obviously

limits the amount of vacancies available to choose from, particularly if the individual was previously unskilled. We interpret our findings as evidence of a strong locking-in effect.

At the end of the paper we try to extract some lessons that may be useful for the future design of social experiments. The problem in Denmark is that - for many types of active labour market policies - social experiments can be useful in order to obtain any kind of impact estimator, due to the mandatory nature of the active labour market policy conducted; everyone who has been unemployed for more than a year has a right and a duty to participate in some kind of activity, like subsidised employment, training, etc. Hence, for any comparison group to exist, we need to be able to have some individuals who are forced not to participate. The social experiment that we investigate in this paper, the first of its kind in Denmark, does provide us with some important insights into how to design future experiments.

Our study is by no means the first to find very small effects of classroom training in Denmark on the subsequent individual unemployment rate.<sup>1</sup> Indeed, Jensen, Pedersen, Smith & Westergaard-Nielsen (1993) and Westergaard-Nielsen (1993) find very small or no effects of participation in labour market (AMU) training on individual unemployment rates. Our study is based on the same data as Langager (1996). He is still the only one to find consistently disappointing effects of AMU training<sup>2</sup>.

Internationally, social experiments have received some attention, with experiments being conducted mainly in the US (see e.g. the review in Heckman, LaLonde & Smith (1999)). However, in Europe very few social experiments have been conducted. The few that have come to our attention have been performed in Norway (Torp, Raaum, Hernæs & Goldstein (1993), Raaum, Torp & Goldstein (1994)), Denmark (Langager (1996)), Sweden (Delander (1978), and Björklund & Regné (1996)), the UK (White & Lakey (1992)), and the Netherlands (Van Den Berg & der Klaauw (2001)). The main findings from experimental evaluations in the U.S are that training impacts on earnings<sup>3</sup> are modest, yet positive for some demographic groups, especially economically disadvantaged women, and to some extent also economically disadvantaged men.<sup>4</sup> On the other hand, youth training is mostly ineffective. It is also evident that within each of these groups, those who benefit the most are the most able in terms of education, and past employment and earnings histories. The only other experiment conducted in Europe (to our knowledge) that analyses training programmes is Torp et al. (1993), and they find moderate, but positive effects on the employment rates of trainees. In Denmark, the only experimental study performed (Langager (1996)) concludes - as mentioned above - that training increases the subsequent unemployment rates of participants for all demographic groups.

Looking at the evidence on the impact of training using non-experimental data, the general finding from the U.S. literature is that there is very wide dispersion in the estimated effects of training on earnings. However, reviewing many of these non-experimental evaluations (and discarding those most likely to produce biased estimates), Heckman et al. (1999) conclude that the evidence is consistent with the findings from the U.S. experimental literature, namely that training significantly increases the earnings of economically disadvantaged women, that training in some cases increases the earnings of economically

---

<sup>1</sup>The type of classroom training studied here is labour market training, called AMU, and will be described briefly in section 2.

<sup>2</sup>See, however, our discussion of his result in the end of section 5.

<sup>3</sup>There is a close correspondence between earnings impact and impacts on employment rates, in the sense that positive impacts on earnings often are associated with positive impacts on employment rates. This is, of course, the case if training affect employment probabilities rather than the hourly wage.

<sup>4</sup>cf. Heckman et al. (1999)

disadvantaged men, and finally, that training has no measurable effect on the earnings of youth. In Europe, the general finding is that training increases employment rates, although the effects are modest and not always significant (see e.g. Bonnal, Fougere & Serandon (1997); Kraus, Puhani & Steiner (1997); Björklund (1993); Björklund (1994); Harkman, Jansson & Tamas (1996); Richardsson & Berg (2002)).

The remainder of this paper is organised as follows: Section 2 presents a very brief summary of the institutions and rules of labour market training in Denmark. In Section 3 the experiment and the data used for this study are presented. Section 4 contains the experimental methodology and impact estimates. We take a closer look at the actual selection process into training in section 5. In section 6 we formulate and estimate a dummy endogenous variable tobit model, and in section 7 various matching estimators and impact estimates obtained from these are presented. Finally, section 8 concludes the analysis, extracts the main lessons for future experiments and points towards potentially fruitful research paths that may be pursued in the future.

## 2 Labour Market Training in Denmark<sup>5</sup>

The purpose of the Danish system of Labour Market Training Programmes (the AMU-programme) is to maintain and improve the qualifications of the work force. In addition, it should assist in facilitating adjustments to structural changes in the labour market (say, shifts in the demand for labour). The programme is large, and it operates on many target groups. The type of training considered in this study is 'competence enhancing training for (mainly) unskilled workers'<sup>6</sup>. Training takes place at one of 24 training sites. In 1993, 118,000 persons undertook training of this type, corresponding to 4.1% of the labour force, and to 45% of all trainees at AMU-sites in 1993. There are approximately 600 different training courses to choose from. A course of this type lasts from 1 to 7 weeks, with an average duration of 2 weeks. These courses are open for employed as well as unemployed workers. When there are more applicants for a given course than the maximum capacity of that course, the selection procedure is such that employed workers are given first priority, unemployed workers with a work contract conditional on participation in the course are given second priority, and other unemployed workers have the lowest priority. Unemployed workers can obtain unemployment insurance benefits while in training.

## 3 The Experiment and the Data Set

The AMU experiment was conducted in 1994. The types of courses under evaluation was restricted to courses directed towards certain industries. As a consequence, 57% of all training participants in the experiment received training in the area of land transportation (mainly certificates for pick-up truck, truck, crane, bus etc.), as opposed to only 30% in the population of participants. The remainder of courses were mainly in the areas of 'metal industry courses' and 'introductory computer courses'<sup>7</sup>.

---

<sup>5</sup>This section draws on Langager (1996).

<sup>6</sup>Other types of training under the AMU programme targeted at different groups of workers and/or firms include competence enhancing training for skilled workers, training directed towards firm specific needs, and longer lasting training primarily for the unemployed.

<sup>7</sup>Langager (1996) investigates whether different courses have different effects, but finds no evidence of this. In this study, we therefore consider all types of training as identical in the sense that they have the same effect on unemployment.

Unemployed individuals who in April 1994 applied for participation in one of the selected AMU-courses with starting date in May and June 1994 were randomly assigned into a treatment and a control group. Treatment group members were then offered the course, while control group members were not. Surveys were conducted in December 1994, and this information was matched with administrative records containing information at the individual level on weekly unemployment status (a degree of unemployment ranging from 0 to 1) and a few other variables. The experiment included 938 persons in total. We had to delete 126 of these due to nonresponse or other missing data problems. This leaves us with 812 persons, of whom 425 were in the treatment group and 387 were in the control group. Of the 425 in the treatment group, 219 (52%) received training while the remaining 206 did not show up. These two subgroups of the treatment group are called the **Treated** and **No-Shows**, respectively. In the control group, 86 (22%) ended up receiving training during May and June 1994 in one of the courses they were supposedly randomised out of. This is the group of **Cross-Overs**, while the last group, those in the control group not receiving treatment in the treatment period, are denoted **Real Controls**. It is thus a highly imperfect experiment, as ideally there should have been 100% treated in the treatment group and 100% real controls in the control group.

The dependent variable in this study is the individual unemployment rate. This variable is available on a weekly basis, but we have chosen to group it. In appendix B, figures 2-5 show kernel density estimates of the dependent variable in pre- and post-treatment periods for various groups in the sample<sup>8</sup>. Figures 6-7 show the mean individual weekly unemployment rate for some of the sub-samples used. The pre-participation period runs from November 1993 until April 1994 both months inclusive, that is, weeks 124-149 in the figures, while the post-participation period runs from July 1994 until December 1994, both months inclusive, that is, weeks 159-184 in the figures. It is evident that individuals in both treatment and control groups are most unemployed in the month of application (and hence selection), April, which is to be expected, it is merely a manifestation of Ashenfelter's dip (or rather peak, in this case), see e.g. Heckman & Smith (1999). After the selection period, there is a regression to the mean process going on, whereby unemployment falls again for both treatment and control groups. Judging from these figures, it may also be noted that there are some notable differences between the subgroups of the treatment and control groups already in the pre-participation period, in the sense that the treated and cross-overs are more unemployed already prior to participation than the no-shows and real controls. This implies that selection is non-random, an issue that will be addressed again in Section 5 and onwards. Note that the dependent variable is potentially both left- and right-censored (at 0 and 1, respectively). For the pre-participation period, about 20 percent of the observations are left-censored, and 12 percent are right-censored, while for the post-participation period, the fractions are 30 percent and 5 percent, respectively, with minor differences between treatment and controls in the post-treatment period.

---

<sup>8</sup>Note that the endpoints represent point probabilities corresponding to the fractions of left- and right-censored observations. The kernel densities are scaled so that they integrate to the fractions of uncensored observations.

## 4 Experimental Methods and Results

### 4.1 Methodology

In this section, we will discuss the experimental estimators applied in this study and the identifying assumptions underlying each of these estimators. The limited nature of the dependent variable leads us to estimate the group means by maximum likelihood in a censored regression type model. The main identifying assumptions relate to the question of whether the objects are aware of participation in the experiment and whether the presence of the experiment will affect their behaviour. An additional problem is introduced by the social nature of the experiment, that is, by the fact that individuals do not comply with their assigned status as treated and controls. This is the problem of no-shows and cross-overs. These terms reflect that there will be some individuals in the treatment group who, even if they applied to participate in training voluntarily, do not show up on the first day of training (no-shows). Equivalently, some of the persons assigned to the control group will end up being in training anyway (cross-overs), either because of ethical problems when there are vacant training slots or because of a very active effort put in by some individuals trying to obtain the desired training. We will modify the simple experimental estimator to take account of this problem, stipulating the relevant identifying assumption for this (non-experimental) estimator.

We begin with a brief outline of some notation, assumptions, and a formulation of the traditional experimental estimator. The objective of the evaluation is to measure the effect or impact of a treatment,  $D$ , on an outcome variable,  $Y$ . Let  $Y_{1i}$  be the person-specific outcome in the presence of treatment, and  $Y_{0i}$  the outcome in the absence of treatment, both defined under normal programme operation. Hence, the person-specific impact of the programme is defined as  $\Delta_i = Y_{1i} - Y_{0i}$ . The fundamental evaluation problem is that we do not observe the same person with both outcomes,  $Y_1$  and  $Y_0$ . It therefore becomes impossible to construct the person-specific impact for anyone by simply looking at the data. Instead, attention usually shifts to constructing means. The usual parameter of interest in the literature is the average effect of "treatment on the treated" defined as<sup>9</sup>

$$\begin{aligned}\Delta &= E[Y_1 - Y_0 | D = 1] \\ &= E[Y_1 | D = 1] - E[Y_0 | D = 1]\end{aligned}\tag{1}$$

Hence, the problem shifts to that of finding the counterfactual  $E[Y_0 | D = 1]$  in (1), which is unobserved but must be constructed in order for the defined impact measure to be identified, i.e. some assumptions are needed to obtain identification. These assumptions depend on the type of data available. We distinguish between experimental data and non-experimental or observational data. In this study we use experimental data (at least the data is called experimental, although its experimental nature will be called into question in this paper).

One reason for conducting social experiments is that they produce simple impact estimates that are easy to calculate and are easily understood and interpreted, even by policy makers:-). The basic idea is the following: Select a group of would-be participants in the programme - for instance applicants who would be admitted into the programme under normal operating conditions;  $D = 1$  - and split them randomly into a treatment group ( $R = 1$ ) and a control group ( $R = 0$ ). The members of the treatment group receive the

---

<sup>9</sup>There exists a variety of impact measures, but since we will not apply those in this study, they are not discussed. See Heckman et al. (1999) for an overview.



treatment, while the control group members do not. Hence an assumption required for the impact estimate produced by our social experiment to correspond to the treatment on the treated parameter (1) is the absence of randomisation bias<sup>10,11</sup>.

Given the traditional experimental assumption that

$$\begin{aligned} E[Y_1|D = 1, R = 1] &= E[Y_1|D = 1] \\ E[Y_0|D = 1, R = 0] &= E[Y_0|D = 1] \end{aligned} \tag{A-1}$$

$\Delta$  is identified and equal to

$$\Delta = E[Y_1|D = 1, R = 1] - E[Y_0|D = 1, R = 0] \tag{2}$$

Assumption (A-1) claims that with experimental data we can use the mean in the control group as an estimator of the unobserved counterfactual in (1). There are various problems associated with conducting experiments. These are ethical problems, problems of randomisation bias, that is, violations of (A-1) above, and finally there is the problem that the experiment is a social rather than a controlled experiment. This means that - since we are dealing with human beings - we can never entirely control the environment or their behaviour, and hence various problems may arise that the researcher had not anticipated. The subsequent solution to these problems is one of the most challenging tasks in the literature on experimental evaluation.

The traditional experimental impact estimator is

$$\hat{\Delta}_1 = \frac{1}{n_1} \sum_{\{i:R=1\}} Y_i - \frac{1}{n_0} \sum_{\{j:R=0\}} Y_j \tag{3}$$

and it consistently estimates the impact in (1) using post-treatment data of the outcome of interest, if (A-1) holds.  $n_1$  and  $n_0$  refer to the size of the treatment and control groups, respectively.

However, in the present case, where the dependent variable is limited, some modifications are needed. Let us assume that (A-1) holds, and that, in addition, the dependent variable is such that

$$\begin{aligned} Y_0^* &= \mu_0 + \varepsilon_0, & \varepsilon_0 | D = 1 &\sim N(0, \sigma_{\varepsilon_0}^2) \\ Y_1^* &= \mu_1 + \varepsilon_1, & \varepsilon_1 | D = 1 &\sim N(0, \sigma_{\varepsilon_1}^2) \end{aligned} \tag{A-N}$$

and for  $D = \{0, 1\}$

$$Y_D = \begin{cases} Y_D^* & \text{if } 0 < Y_D^* < 1 \\ 0 & \text{if } Y_D^* \leq 0 \\ 1 & \text{if } Y_D^* \geq 1 \end{cases}$$

The observed variables are  $Y_0$  or  $Y_1$ , while the underlying latent variables are  $Y_0^*$  and  $Y_1^*$ . Given (A-1) and the additional distributional assumption (A-N) made above, we can

---

<sup>10</sup>Note that the common effect assumption,  $\Delta = \Delta_i$ , can be used as a substitute for this no-randomisation bias assumption. Under the common effect assumption, it does not matter whether the same persons are treated under random assignment, as everyone will experience the same impact in any case.

<sup>11</sup>Besides this we also need the assumption of no "general equilibrium effects" in order for our analysis to be but internally valid. Hence we would have to assume that persons outside our experimental treatment and control groups are not affected by the treatment. This is known as the "stable unit treatment value assumption" (SUTVA) in the statistics literature.

focus on the marginal effect of training on non-limit observations, that is, the effect of training on  $Y^*$ , rather than  $Y$ . This comes, however, at the cost of having to assume a common treatment effect  $\Delta \equiv \mu_1 - \mu_0$ . The parameters of the model are estimated by maximum likelihood. Denote the estimate of the effect of training obtained from this model  $\hat{\Delta}_{ML1}$ . We then have

$$\hat{\Delta}_{ML1} = \hat{\mu}_1 - \hat{\mu}_0 \quad (4)$$

The impact estimate can be made conditional on explanatory variables, of course, by selecting subsamples of individuals, or by specifying fully interacted tobit models, the difference being the restriction on the variance implied by the fully interacted model. In this paper, we consider only mean effects.

A social experiment is not completely under the control of the researcher. In particular, there is a risk of non-compliance. Let  $T$  indicate whether an individual is actually receiving treatment. We may then split the treatment and control groups into the following categories:

<b>Treated:</b>	Treatment group members receiving treatment $D = 1, R = 1, T = 1$
<b>No-Shows:</b>	Treatment group members not receiving treatment $D = 1, R = 1, T = 0$
<b>Real Controls:</b>	Control group members not receiving treatment $D = 1, R = 0, T = 0$
<b>Cross-Overs:</b>	Control group members receiving treatment $D = 1, R = 0, T = 1$

In this case, the estimator (3) or (4) does not estimate the effect of treatment on the treated, but rather the effect of the intention to treat (Heckman et al. (1999)). Within group selection may not be random. Let  $p$  denote the fraction of treated in the treatment group, and let  $q$  denote the fraction of treated in the control group (cross-overs). Under the following assumption

$$E[Y_1 - Y_0 | D = 1, R = 1, T = 1] = E[Y_1 - Y_0 | D = 1, R = 0, T = 1] \quad (\mathbf{A-2})$$

which is a non-experimental type of assumption about the selection process, the following estimator is consistent<sup>12</sup>

$$\Delta = \frac{E[Y | D = 1, R = 1] - E[Y | D = 1, R = 0]}{p - q} \quad (5)$$

Assumption **(A-2)** implies that the expected gain from participation is the same among those treated in the control group as it is for those in the treatment group who end up receiving the training.

This can be estimated in the following way

$$\hat{\Delta}_2 = \frac{\hat{\Delta}_1}{p - q} \quad (6)$$

for the case of uncensored data<sup>13</sup>.

<sup>12</sup>We could also assume a common effect. See Heckman et al. (1999)

<sup>13</sup>When estimating the impact by maximum likelihood techniques we can of course not correct the inconsistency introduced in the first step by simply rescaling the parameter.

## 4.2 Results

In Table 1, we present results obtained by applying the three different experimental estimators to the data. The first column of results comes from treating data as uncensored. The second column of results comes from estimating the mean separately for each group in the sample allowing for different variances, that is, performing a separate maximum likelihood estimation for each group. Remember that due to the nature of our data,  $\hat{\Delta}_1$  and  $\hat{\Delta}_2$  are two different parameters and as such do not estimate the same thing; the first now being an *intention* to treat, and the second the actual treatment effect on those taking the treatment.

**TABLE 1**  
EXPERIMENTAL ESTIMATES  
Unemployment Rate<sup>a</sup>  
Full Sample, 425 in Treatment group and 387 in Control group

	Raw Means	ML Means
$\hat{\Delta}_1$	0.02 (0.02)	0.03 (0.04)
$\hat{\Delta}_2$	0.07 (0.08)	

<sup>a</sup> Std. errs. appear in parentheses.

As it is evident, all three impact estimates are positive, implying that labour market training increases the subsequent unemployment rate of participants relative to non-participants. The traditional experimental estimator  $\hat{\Delta}_1$  and its 'ML' counterpart produce low and insignificant impact estimates. According to these estimators, the intention of receiving training increases individual unemployment by 2-3.5 percentage points. Acknowledging the presence of no-shows and cross-overs, implying that the difference in treatment intensity between the treatment and control groups is only 0.3 (where ideally it should have been 1), we move on to estimator  $\hat{\Delta}_2$ . In order for this estimator to be meaningful we have had to resort to making assumption **(A-2)**. The estimator implies much larger - 7 percentage point increases in the individual unemployment rate - but still insignificant training impacts. Nevertheless, this result suggests an important point, namely that the presence of no-shows and cross-overs can seriously bias the parameter estimates if we do not take properly account of them.

The estimator used by Langager (1996) is a 'raw mean' version advocated for by making the following assumption

$$\begin{aligned} & E [Y_{0,t} - Y_{0,s} | D = 1, R = 1, T = 1] && \textbf{(A-3)} \\ = & E [Y_{0,t} - Y_{0,s} | D = 1, R = 0, T = 0] \end{aligned}$$

where  $s$  and  $t$  refer to the pre- and post-treatment period, respectively. This alternative assumption states that even though there is a non-random component in the selection process into training among both groups, this component is permanent and therefore it can be "differenced" out. This leads to a potential problem known as Ashenfelter's Dip. The problem is that individuals applying for training have extraordinarily low earnings (here: high unemployment rates) and over time, a regression towards the mean will take place, whether or not the individual is trained, see for example figures 6 and 7 where evidence of the Dip (or peak, in this case) is strongly present. The suggested solution to

that problem is to employ a difference-in-differences estimator to the group of treated and real controls

$$\begin{aligned} \Delta_3 = & E [Y_{1,t} - Y_{0,s} | D = 1, R = 1, T = 1] \\ & - E [Y_{0,t} - Y_{0,s} | D = 1, R = 0, T = 0] \end{aligned} \quad (7)$$

and may (in practice) be estimated by

$$\widehat{\Delta}_3 = \frac{1}{n_{11}} \sum_{\{i:R=1,T=1\}} (Y_{i,t} - Y_{i,s}) - \frac{1}{n_{00}} \sum_{\{j:R=0,T=0\}} (Y_{j,t} - Y_{j,s}) \quad (8)$$

in the case of an uncensored continuous dependent variable, where  $n_{11}$  is the number of treated, and  $n_{00}$  is the number of real controls. This estimator compares only the treated in the treatment group with the real controls in the control group. The assumption 7 implies that - although individuals differ in unemployment levels - the change over time in their unemployment history would be similar, if it were not for the training. Any significant difference-in-differences is then assigned to the training programme. However, even though the group of real controls is also expected to experience a regression towards the mean, there is nothing that ensures that this regression is of the same order of magnitude. In fact, it is likely that real controls may experience an even more powerful regression towards the mean, since one (very likely) reason for not participating in training is that the person has found employment. The underlying assumption is thus that the regression to the mean is of the same size for both groups in the absence of training. The estimators  $\widehat{\Delta}_3$  produce impact estimates from 4.7 to 7.7 percentage point increases, although none of them are significantly different from zero.<sup>14</sup> Langager (1996) did find a significant coefficient using this estimator, but the reason for that was his different choice of pre- and post-treatment period; Langager chose as the pre-treatment period a 15 month period immediately before training, and as a post-training period an 8 month period immediately after training. In this paper, we used a symmetric definition (in terms of the period length) of 6 month immediately before and after the training months (May and June). The reason for choosing only a six month window<sup>15</sup> is that the two last post-treatment months contain erroneous data. The choice of relevant pre-treatment window and the resulting differences in the estimates reveals a potential weakness of the estimator (7).

## 5 Probability of Programme Participation

Our next step is to have a look at the condition that treated and real controls are randomly selected from within the treatment and control groups, respectively. As noted above, we face the dual problem of persons in the treatment group who fail to show up for treatment (no-shows) and persons in the control group who receive training anyway (cross-overs). Taking a closer look at figure 7, it does indeed appear to be the case that there is non-random selection, in the sense that the group of treated are on average more unemployed than the real controls already in the pre-training period. As it was also hypothesised in Section 4, the employed workers may be over-represented among the real controls and no-shows. Particularly, since this selection process takes place during April, May, and

---

<sup>14</sup>Tightening the assumption 7 slightly, one can use also the no-shows and the cross-overs. This estimator also produces insignificant estimates around 5-7 percentage points.

<sup>15</sup>We have experimented with observation windows as short as one month, see the next sections.

June of 1994, the dynamics of unemployment may be very different for the four groups, which is exactly the reason why assumption (A-3) is most likely not to be satisfied. Hence, we would expect the difference-in-differences estimator to be biased towards the finding that training is bad. Assumption (A-2) implies that the selection processes in the treatment and control groups are identical. This means that even though a larger fraction in the treatment group ends up as participants than is the case in the control group, we still assume that the expected returns to training are identical on average in the two participant groups (treated and cross-overs). Suppose that they are not, and that - since the cross-overs are a smaller fraction of the control group than the treated are of the treatment group, and since it is more difficult for control group members to obtain training - the expected impact of training is larger on average among the cross-overs than among the treated. Then we will find again that training is bad.

Taking a closer look at the differences between those who actually enroll in training and those who do not, for the two different experimental groups, that is, for the groups of  $R = 1$  and  $R = 0$  separately, the following differences become apparent from table 2: Cross-Overs appear to be slightly younger than Real Controls which is also intuitively appealing from a human capital view point. The younger the participant, the longer is the time period in which to reap the payoffs from the investment in training. However, the picture is not so clear cut for the treatment group, where the difference between Treated and No-Shows is only half a year. More males seem to show up for the training than females in the treatment group, whereas the opposite is the case in the control group. The group actually receiving treatment in the treatment group has slightly more labour market experience in accordance with their older age, whereas the Cross-Overs are less experienced than Real Controls. There are no particular differences in their previous education, nor their health status. Cohabitation does not seem to influence the participation decision for any of the two groups, but the presence of a young child may lower the effort devoted to obtaining training among the controls. A substantially larger fraction of Cross-Overs than Real Controls are not members of an unemployment insurance fund, that is, they do not receive unemployment insurance benefits.

The pre-programme unemployment peak is clearly evident as a potential selection mechanism. Both Treated and Cross-Overs are experiencing a much higher degree of unemployment in the month leading up to randomisation. The Treated did apply for the course expecting positive effects on their labour market attachment more systematically than the group of No-Shows, whereas the opposite is observed among the controls. The importance of previous labour market dynamics in the context of programme participation is also evident when looking at the number of previous jobs leading up to the time of application. The fraction of participants without a real job in 1994 is higher for both treated and cross overs than it is for their experimental counterparts. The fraction having received encouragement to participate by either a former employer or by a caseworker are not surprisingly higher for both participation groups, whereas more Cross-Overs felt forced to participate than Real Controls. The fraction of people having been employed in a real job in the private sector as their latest job is also higher for both the Treated and Cross-Overs than their counterparts. Looking at the training sites it seems that compliance problems among the control group were larger in Randers than at the other training sites. Nearly twice as many as in the balanced case come from Randers among the Cross-Overs, indicating that the protocol of the experiment may have been discarded to a larger extent at this site than at the remaining four sites, potentially due to problems in filling out the open slots at the courses.

We now proceed to conducting a more thorough investigation of the actual participation decision among the groups. This is done by running a simple probit model for the pooled data. The probability of programme participation,  $P(T = 1|\mathbf{X})$ , plays a central role in our two non-experimental estimators implemented below. In our setting participation means participation conditional on application.

Previous international studies (see Heckman et al. (1999) and Heckman & Smith (1999)) have shown that a key predictor of participation is recent labour market dynamics/transitions. So persons recently entering unemployment either from outside the labour force or from previous employment are most likely to seek participation in programmes. Our experimental sample is constructed in a way such that all persons are unemployed at the time of application. In spite of this homogeneity of the sample, even small variations in the labour market histories (number of jobs leading up to time of application and the degree of unemployment prior to this) are still found to have significant predictive power of participation. Table 3 below clearly demonstrates this point. As indicated both by figure 7 and by the raw means the pre-programme unemployment rate does significantly influence the decision to participate. The more unemployed the higher is the probability of receiving the course indicating both the lower opportunity cost of participation (expected foregone earnings are lower) and that these courses may be used among the unemployed to regain some attachment to the labour market. At the same time we see that some attachment to the labour market in the year prior to participation has a positive impact compared to being out of employment, whereas the situation is opposite for the time running up to application; having had a job significantly reduces the probability of participation. The additional variables included are based on what is found in the literature to be key predictors (see Heckman et al. (1999)).

Despite the predictions from human capital theory there are no significant differences in the age groups among the participants and non-participants. But there are significant differences between men and women in the selection process as indicated by the discussion of the means above. Men have a higher probability of participation than women, but for women initially randomised out, there is a significant effect on participation. It could be the case that women seek different courses than men, and that these courses are harder to fill. Hence, ethical considerations may force caseworkers to assign women initially randomised out to these empty slots. However, nothing in the data suggests that women randomised out participate in courses that are in any way different from the other courses. Instead, we conjecture that there might be an "empathy" effect from the caseworkers in charge of running the experiment.<sup>16</sup>

No significant differences emerge between the educational categories in the participation probabilities. Counter-intuitively, having expectations of an employment gain by taking the course upon applying does not significantly raise the probability of actually showing up for the course a month later. The estimate on this indicator is found not to be significant at any conventional level. This indicates either that another stronger motive is determining actual participation or that people by some way review their expectations about the potential gains from the course in the time period between application and course start.

---

<sup>16</sup>They know that the woman in front of them is supposed to be turned down but overrule the experimental protocol and assigns her to the course anyway.

**TABLE 2**  
DESCRIPTIVE STATISTICS FOR THE DIFFERENT EXPERIMENTAL GROUPS<sup>a</sup>

Variables <sup>b</sup>	All	Treated	No-Shows	Real Controls	Cross-Overs
Age	31.55 (9.42)	31.70 (10.13)	31.28 (9.47)	31.91 (9.17)	30.56 (8.33)
Male	0.76	0.81	0.67	0.82	0.66
Immigrant	0.05	0.06	0.03	0.04	0.06
Self-reported labour market experience, years	11.24 (10.05)	11.42 (10.66)	10.87 (10.21)	11.83 (9.84)	9.59 (8.71)
Fewer than 9 years of schooling	0.13	0.12	0.15	0.12	0.10
Highschool education	0.04	0.07	0.04	0.02	0.01
Vocational education	0.37	0.35	0.38	0.38	0.31
Short further education	0.03	0.01	0.02	0.05	0.05
College education and beyond	0.02	0.03	0.02	0.02	0.01
Self-reported health status, medium	0.10	0.08	0.09	0.12	0.14
Self-reported health status, poor	0.03	0.02	0.03	0.03	0.01
Living with a partner	0.63	0.63	0.65	0.61	0.63
Partner unemployed during last year	0.24	0.26	0.24	0.23	0.24
Child age less than 6 present in household	0.24	0.22	0.24	0.24	0.26
No Unemployment Insurance	0.11	0.10	0.10	0.10	0.20
UI in Metal	0.07	0.06	0.07	0.07	0.13
UI in KAD (female)	0.06	0.04	0.11	0.05	0.05
UI in Construction	0.04	0.01	0.01	0.07	0.05
UI other place	0.27	0.31	0.26	0.25	0.23
Unempl. rate prior to randomisation	0.52	0.60	0.41	0.49	0.64
Applied based on expected employment gain	0.72	0.78	0.71	0.71	0.66
Encouraged to participate	0.13	0.16	0.09	0.10	0.20
Forced to participate	0.02	0.01	0.02	0.01	0.05
One job spell in 1993	0.57	0.57	0.54	0.57	0.64
Two job spells in 1993	0.14	0.14	0.16	0.15	0.12
More than two job spells in 1993	0.03	0.04	0.03	0.03	0.05
One job spell in 1994 prior to RA	0.43	0.36	0.54	0.45	0.33
Two or more job spells in 1994 prior to RA	0.06	0.04	0.11	0.06	0.02
Last job in public sector	0.04	0.05	0.03	0.03	0.06
Last job in private sector	0.36	0.44	0.24	0.37	0.42
Aabenraa	0.18	0.15	0.21	0.17	0.17
Aalborg	0.23	0.23	0.21	0.25	0.20
Odense	0.20	0.22	0.19	0.21	0.13
Randers	0.23	0.23	0.22	0.20	0.36
<b>Sample size</b>	<b>812</b>	<b>219</b>	<b>206</b>	<b>301</b>	<b>86</b>

<sup>a</sup> Standard deviations shown in parentheses.

<sup>b</sup> Omitted educational group is 9 or 10 years of schooling, omitted health category is good / excellent, omitted UI category is unskilled (SID), omitted place of last employment is job created as a means of fighting unemployment, and omitted training site is Slagelse.

**TABLE 3**

COEFFICIENT ESTIMATES AND ASY. STD. ERR. FROM PARTICIPATION PROBIT  
 Dep. variable: 1 for Treated and Cross-Overs, 0 for No-Shows and Real Controls, Full Sample

Variables <sup>a</sup>	Coeff	Asy Std Error
Intercept	<b>-2.50</b>	0.65
Aabenraa	0.11	0.19
Aalborg	0.02	0.16
Odense	0.19	0.17
Randers	<b>0.40</b>	0.17
Male	<b>0.41</b>	0.18
Woman with 'R=0'- status	<b>0.95</b>	0.23
Immigrant	0.37	0.25
Age below 25	0.55	0.60
Age 25 to 29	0.59	0.58
Age 30 to 39	0.46	0.54
Age 40 to 49	0.29	0.49
Age 50 to 54	0.34	0.49
Fewer than 9 years of schooling	-0.24	0.18
High School education	0.27	0.25
Vocational education	-0.21	0.13
Short further education	-0.32	0.31
College education and beyond	0.15	0.34
Living with a partner	0.12	0.13
Partner unemployed during last year	-0.01	0.13
Child age less than 6 present in household	-0.06	0.13
Self-reported labour market experience, years	0.01	0.01
One job spell in 1993	<b>0.29</b>	0.12
Two job spells in 1993	0.18	0.17
More than two job spells in 1993	<b>0.72</b>	0.34
One job spell in 1994 prior to RA	<b>-0.44</b>	0.13
More than one job in 1994 prior to RA	<b>-0.85</b>	0.27
Last job in public sector	0.15	0.27
Last job in private sector	0.11	0.12
Self-reported health status, medium	-0.00	0.17
Self-reported health status, poor	-0.42	0.30
Applied based on expected employment gain	0.09	0.12
Encouraged to participate	<b>0.55</b>	0.15
Forced to participate	0.62	0.40
Individual unempl. rate in month prior to RA	<b>0.38</b>	0.13
No Unemployment Insurance	0.16	0.19
UI in Metal	0.20	0.21
UI in KAD (female)	-0.42	0.25
UI in Construction	0.02	0.29
UI other place	0.02	0.12
R	<b>1.17</b>	0.12

<sup>a</sup>The omitted training center is Slagelse; omitted age group is above 54; omitted schooling category is 9-10 years; omitted health status is very good/good; omitted caseworker/employer assesment is not encouraged / forced; omitted UI category is SID (unskilled workers)



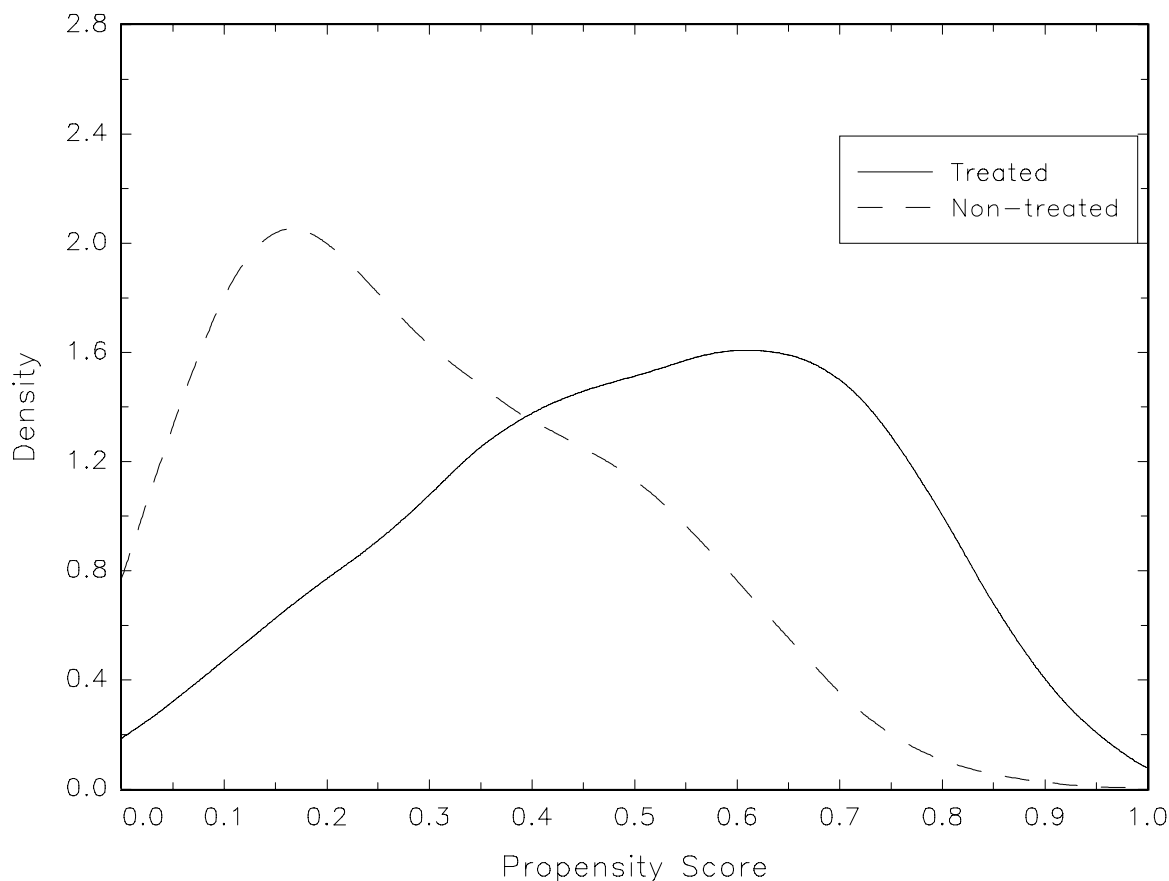


Figure 1: Nonparametric estimates of propensity densities (Biweight Kernel), probit  $P(X)$  model; bandwidth using Silverman (1986) rule of thumb.

Being recommended by the previous employer or a caseworker to participate increases significantly the subsequent probability of participation for both groups whereas feeling forced to participate does not significantly increase the likelihood of participation. Finally, the training site 'Randers' comes out with a significantly higher coefficient on the probability of receiving training than the other training sites, a phenomenon not accounted for in the documentation we have available.

Figure 1 presents the kernel densities of the estimated  $P(T = 1|\mathbf{X}) = P(\mathbf{X})$  in the  $\{T = 0\}$  and  $\{T = 1\}$  groups. The distributions obtained for  $P(\mathbf{X})$  using a variety of different covariates all produce the findings of the figure shown here. The two groups are very close to have a common support of  $P(\mathbf{X})$ . This comes of course as no surprise given the homogeneous sample of the experiment. But as is clearly evident from the picture the shape of the distributions differ over the region of common support. The treated have a mode of the distribution at 0.65 whereas many non-treated have a low estimated probability of participating. The main reason for the distinct difference in the distributions of the propensity scores is naturally the presence of the randomisation indicator as a key predictor of participation.

This discussion does not necessarily alter our preliminary conclusion that training has no effect, but it might. This question is analyzed in the next two sections, where we

look at non-experimental estimators to deal with the issue of selection bias. We employ two different types of estimators. The first, a dummy endogenous variables model, takes account of the self-selection based on unobservables and of the censored nature of our dependent variable. However, it does this at a cost of assuming a common treatment effect and a potential misspecification in terms of distributional assumptions. The second method, matching, is nonparametric in nature and allows for the treatment effect to vary across individuals but the selection is based on observables or to some extent on temporally persistent unobservable factors, and we do not address the censoring of the dependent variable.

## 6 A Dummy Endogenous Variable Model

Assume that we may pool the treatment and control group members. We can then think of the following model

$$\begin{aligned} t_i^* &= \mathbf{z}'_i \boldsymbol{\gamma}_1 + r_i \gamma_2 + \nu_i \\ t_i &= \begin{cases} 1 & \text{if } t_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \\ y_i^* &= \mathbf{x}'_i \boldsymbol{\beta} + \Delta t_i + \varepsilon_i \end{aligned} \quad (9)$$

where  $R$  is included to capture the fact that selection into training may be more difficult for control group members than for treatment group members.

In the equation for the outcome of interest,  $Y$ ,  $R$  has no explanatory power once the participation indicator is included, since randomisation is - random. Hence, the treatment group indicator is an ideal instrument for identifying the parameters of a model with a common treatment effect, as it is by construction uncorrelated with the outcome variable,  $Y$ , and strongly correlated with the receipt of training. The error terms are assumed to be bivariate normally distributed, with mean 0 and covariance matrix

$$\text{cov}(\nu, \varepsilon) = \begin{Bmatrix} 1 & \rho\sigma_\varepsilon \\ \rho\sigma_\varepsilon & \sigma_\varepsilon^2 \end{Bmatrix}$$

The likelihood function is (through a modification of the likelihood function derived by Heckman (1978), for linear models)

$$\begin{aligned} \mathcal{L} &= \prod_{\{i:0 < y_i^* < 1\}} f(y_i^*, t_i) \cdot \prod_{\{i:y_i^* \leq 0\}} \Pr(Y_i^* \leq 0, t_i) \cdot \prod_{\{i:y_i^* \geq 1\}} \Pr(Y_i^* \geq 1, t_i) \\ &= \prod_{\{i:0 < y_i^* < 1\}} \left\{ f_\varepsilon(Y_{i\tau} - \mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i) \cdot [1 - F_{\nu|\varepsilon}(-\mathbf{z}'_i \boldsymbol{\gamma}_1 - r_i \gamma_2 | Y_{i\tau} - \mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i)]^{t_i} \right. \\ &\quad \left. \cdot F_{\nu|\varepsilon}(-\mathbf{z}'_i \boldsymbol{\gamma}_1 - r_i \gamma_2 | Y_{i\tau} - \mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i)^{1-t_i} \right\} \\ &\quad \prod_{\{i:y_i^* \leq 0\}} \left\{ \frac{F_{\nu,\varepsilon}(-\mathbf{z}'_i \boldsymbol{\gamma}_1 - r_i \gamma_2, -\mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i)^{1-t_i}}{[F_\varepsilon(-\mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i) - F_{\nu,\varepsilon}(-\mathbf{z}'_i \boldsymbol{\gamma}_1 - r_i \gamma_2, -\mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i)]^{t_i}} \right\} \\ &\quad \prod_{\{i:y_i^* \geq 1\}} \left\{ \frac{[F_\nu(-\mathbf{z}'_i \boldsymbol{\gamma}_1 - r_i \gamma_2) - F_{\nu,\varepsilon}(-\mathbf{z}'_i \boldsymbol{\gamma}_1 - r_i \gamma_2, 1 - \mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i)]^{1-t_i}}{\left[ \frac{\{1 - F_\varepsilon(1 - \mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i)\}}{\{F_\nu(-\mathbf{z}'_i \boldsymbol{\gamma}_1 - r_i \gamma_2) - F_{\nu,\varepsilon}(-\mathbf{z}'_i \boldsymbol{\gamma}_1 - r_i \gamma_2, 1 - \mathbf{x}'_i \boldsymbol{\beta} - \Delta t_i)\}} \right]^{t_i}} \right\} \end{aligned} \quad (10)$$

where  $f_\varepsilon$  is a normal density with mean 0 and variance  $\sigma_\varepsilon^2$ ,  $F_\varepsilon$  is the corresponding cdf,  $F_{\nu|\varepsilon}$  is a normal cdf with mean  $\rho\varepsilon/\sigma_\varepsilon$  and variance  $(1 - \rho^2)$ ,  $F_\nu$  is standard normal cdf, and  $F_{\nu,\varepsilon}$  is the bivariate normal cdf specified above.

The full estimation results of the parameters of model (9) are available on request. The variables used are the same as the variables presented in tables 2 and 3. In Table 4 we present the main parameters of interest.

**TABLE 4.**

COEFFICIENT ESTIMATES AND ASY. STD. ERR. FROM TOBIT SELECTION MODEL

Dep. variables: 1 for Treated and Cross-Overs, 0 for No-Shows and Real Controls, Unemployment Rate

Full Sample, 305 Treated and Cross-Overs and 507 No-Shows and Real Controls

	Tobit Model		Linear Model	
	Coeff.	Std.err.	Coeff.	Std.err.
$\gamma_2$	<b>1.17</b>	0.13	<b>1.17</b>	0.13
$\Delta$	<b>0.20</b>	0.10	0.10	0.06
$\rho$	-0.20	0.15	-0.13	0.15

For reasons of comparability with the estimators obtained in the next section, we also report the parameters from the estimation of a linear endogenous variables model. It is seen that the randomisation variable - as expected - has a strong positive influence on the probability of receiving training. In addition, we find that the estimated correlation parameter,  $\rho$ , is negative (implying negative selection bias). This estimate is however far from being statistically significant. We find a very strong positive intramarginal effect of training on the individual unemployment rate, now being significantly different from zero. This result is thus stronger than the results obtained using the traditional experimental estimator, even when it was corrected for noncompliance. The impact estimator obtained in the linear model, however, is not statistically significant. In the next section we will check the robustness of our results with the matching estimator, and we will investigate the evolution of the impact estimate over the six month post-participation period.

## 7 Matching

Matching is based on the assumption that conditioning on observables,  $\mathbf{X}$ , eliminates the selective differences between programme participants and non-participants. More precisely, the method of matching assumes that the econometrician has access to conditioning variables sufficiently rich such that the counterfactual outcome distribution of the participants is the same as the observed outcome distribution of non-participants. By conditioning on the covariates at our disposal, we will thus be capable of balancing the bias coming from the self-selection into our programme.

In focusing on (1) we make the following additional assumption<sup>17,18</sup>

$$E[Y_0 | \mathbf{X}, T = 1] = E[Y_0 | \mathbf{X}, T = 0] = E[Y_0 | \mathbf{X}]. \quad (\mathbf{A-4})$$

<sup>17</sup>Notice that these assumptions are generally weaker than the usual "strong ignorability" assumptions of Rosenbaum & Rubin (1983), which are often both invoked in the literature of matching even though they are overly strong given that the parameter of interest be the treatment on the treated.

<sup>18</sup>Note that in performing the matches we do not take the censored nature of our data into account and that in what follows below conditioning on  $D = 1$  is left implicit.

In order to be able to utilise **(A-4)** it is necessary to make sure that there is a non-participant analogue to each participant, i.e.,

$$\Pr(T = 1 | \mathbf{X}) < 1. \quad (\mathbf{A-5})$$

Notice again that we do not *not* assume away selection bias. Instead we rebalance the bias in much the same way as the experiment initially was set out to do.

We do not want to assume any functional form of the outcome equation as opposed to for example Barnow, Cain & Goldberger (1980). We are therefore potentially faced with the nonparametric curse of dimensionality due to our rich survey and register data. A way to circumvent the curse of dimensionality without imposing arbitrary assumptions is based on the results in Rosenbaum & Rubin (1983). Here the focus is shifted from the set of covariates to the probability of programme participation,  $P(\mathbf{X}) = \Pr(T = 1 | \mathbf{X})$ . As long as **(A-4)** and **(A-5)** hold,

$$E[Y_0 | P(\mathbf{X}), T = 1] = E[Y_0 | P(\mathbf{X}), T = 0] \quad (\mathbf{A-6})$$

over the common support<sup>19</sup>,  $S_P = \text{Supp}(P(\mathbf{X}) | T = 1) \cap \text{Supp}(P(\mathbf{X}) | T = 0)$ . This new conditioning variable,  $P(\mathbf{X})$ , changes our conditional mean assumption into **(A-6)** which together with  $P(\mathbf{X}) < 1$  are the only conditions required to justify propensity score matching to estimate the mean impact on the treated, see Heckman, Ichimura & Todd (1998).

Our different matching estimators implemented below all take on the following typical form

$$\Delta = \frac{1}{m_1} \sum_{i \in I_1 \cap S_p} \left( Y_{1i} - \sum_{j \in I_0} W(i, j) Y_{0j} \right), \quad (11)$$

where  $I_1$  denotes the set of treated and cross-overs,  $I_0$  the set of no-shows and real controls, and  $m_1$  denotes the number of persons in the set  $I_1 \cap S_p$ . Notice how the match for each participant  $i \in I_1 \cap S_p$  is constructed as a weighted average over the outcomes of non-participants, where the weights,  $W(i, j)$ , are constructed such that they depend on the distance between  $P_i$  and  $P_j$ , where  $P_i \equiv P(\mathbf{X}_i)$ . The different matching estimators we implement below differ in how the weights are constructed.

**Nearest Neighbour matching**, which is the traditional, pairwise version of matching, sets

$$W(i, j) = \mathbf{1}_{j = \arg \min_{k \in I_0} \|P_i - P_k\|},$$

where  $\mathbf{1}_{(\cdot)}$  is the indicator function. We thus use the single non-participant with the value of  $P_j$  closest to  $P_i$  as a match. We implement this both with replacement of the non-participant and with a random sampling scheme (of the participants) without replacement of the matched non-participants. The former procedure may potentially inflate the variance due to the risk of using the same non-participant more than once. The latter trades reduced variance resulting from using the same non-participant in the counterfactual mean only once for an increased bias due to the relatively small pool of non-participants which will result in, on average, poorer matches. To avoid this risk of running poor matches we implement the common support requirement by using **caliper matching** (Cochran & Rubin (1973)). The simple idea here is to impose some pre-specified tolerance on the

<sup>19</sup>See below for ways of constructing this set.

maximum distance  $\|P_i - P_j\|$  allowed. This means that a match for a treated person  $i$  is found only if  $\|P_i - P_j\| < \varepsilon$  for some  $j \in I_0$ , and  $\varepsilon$  is a pre-specified tolerance. Treated persons for whom no matches can be found within the caliper are simply excluded from the analysis. Hence, by performing caliper matching, we have a simple way of imposing a common support condition. Of course, we are to set the caliper and a priori there is no way of knowing what choice of tolerance level to set<sup>20</sup>.

The estimators above use at most one person from the comparison group in constructing a match for the programme participant. **Kernel matching** and **Local Linear Regression matching** are nonparametric matching estimators that construct matches for each programme participant using kernel weighted averages over multiple non-participants. Relative to the simple nearest Neighbour matching, by using kernel techniques we reduce the variance of our matching estimates by making use of information from additional non-participants. As a cost, we introduce small sample bias because of the increased distance between participants and matched non-participants, measured in terms of their probabilities of programme participation. The weights used in the kernel matching estimates are

$$W(i, j) = \frac{K\left(\frac{P_i - P_j}{h_n}\right)}{\sum_{k \in I_0} K\left(\frac{P_i - P_k}{h_n}\right)} = \frac{K_{ij}}{\sum_{k \in I_0} K_{ik}},$$

where  $K$  is a kernel function and  $h_n$  is a bandwidth. Similarly, the weight for local linear matching is

$$W(i, j) = \frac{K_{ij} \sum_{k \in I_0} K_{ik} (P_k - P_i)^2 - [K_{ij} (P_j - P_i)] [\sum_{k \in I_0} K_{ik} (P_k - P_i)]}{\sum_{j \in I_0} K_{ij} \sum_{k \in I_0} K_{ik} (P_k - P_i)^2 - (\sum_{k \in I_0} K_{ik} (P_k - P_i))^2}.$$

We use this latter local linear matching in the analyses below instead of the simpler kernel matching because of its desirable statistical properties (see Heckman, Ichimura & Todd (1997) for details). Specifically, one of the advantages of the local linear estimator over simple kernel weighting is that the local linear estimator will converge at a faster rate at boundary points, and, as was depicted in figure 1, both groups have values of  $P(\mathbf{X})$  very close to the boundary value at zero, and there are values of  $P(\mathbf{X})$  at the boundary point of 1 for the participants. Hence, we will be able to put more confidence in the local linear version of matching compared to kernel matching. Due to support considerations (see Heckman, Ichimura, Smith & Todd (1998) for practical details) we use as kernel function,  $K$ , the biweight kernel given by

$$K(s) = \begin{cases} \frac{15}{16}(1 - s^2)^2 & \text{for } |s| < 1 \\ 0 & \text{otherwise} \end{cases}$$

and a bandwidth following Silverman (1986) "rule of thumb".

Since we have access to panel data a natural thing to do is to "weaken" the cross sectional conditional mean assumption (**A-6**) above to an assumption necessary for Difference-in-Difference (D-in-D) matching strategies. With the data at hand we are able to construct

---

<sup>20</sup>Dehejia & Wahba (1999) implement caliper matching in a somewhat different way than we do. They assign equal weights to all non-treated persons within the caliper and should there be no comparison group members within it, they use the nearest single neighbor outside the caliper instead of dropping the participant from the analysis as we do.

matching estimates of the counterfactual that are robust to temporally persistent separable components of bias between the participants and nonparticipants. Hence, we assume that the matching estimator is subject to the same selection bias in the affected and unaffected periods. Again, we let  $s$  represent a time period before the programme and  $t$  a time period after the programme. We then compare the conditional before-after unemployment rates of actual programme participants with those of the non-participants. Formally, we assume

$$E[Y_{0,t} - Y_{0,s} | P(\mathbf{X}), T = 1] = E[Y_{0,t} - Y_{0,s} | P(\mathbf{X}), T = 0]. \quad (\mathbf{A-7})$$

This identifying assumption states that conditional on  $P(\mathbf{X})$ , the biases are the same on average in different time periods before and after the period of participation. Thus, differencing the differences for participants and non-participants will eliminate this bias. Hence, by having access to the same kind of data for both groups in a panel setting we are able to allow selection to be based on time-invariant unobservables (see Heckman, Ichimura, Smith & Todd (1998) for discussion).<sup>21</sup>

The results<sup>22</sup> from the cross sectional matching strategies are presented in table 5<sup>23,24</sup>. As is seen from the bottom row of the table, the impact estimates are slightly smaller than the one found in the previous section for the linear model, which is the relevant comparison, since we focus on the effect on  $Y$  here and not on  $Y^*$ , the marginal effect on non-limit observations. It is also seen that only one of the three different matching strategies produces a significant impact when we consider the whole period. Remember that some of the variation in the calculated impacts of the different matching strategies is simply coming from the different ways of implementing the common support requirement. Hence, the different estimates are for different subpopulations of the treated.

It seems almost implausible that participation in a two week training course, which provides the participant with formal qualifications, such as for instance a bus certificate, should increase subsequent unemployment relative to non-participation. However, the effect may simply be a locking-in effect; while in training, trainees have much lower search intensity and hence leave unemployment at a lower rate than non-trainees. Subsequently, even if the hazard rates of trainees are larger than those of non-trainees, we may find that the average rate of unemployment is larger in the sample of non-trainees than in the sample of trainees. However, this difference should disappear over time, if there is only a locking-in effect. This question is addressed in the table as well, where we have estimated the impact of training month by month in the period following participation. The dynamics of the effect of participation come out more clearly here: Ranging from 0.07 to 0.10 with two of the estimators significantly different from zero, the effect of participation generally declines during the autumn. Hence, since the impact of training is large in the first month

---

<sup>21</sup>As we pointed out in section 3, this type of estimator is potentially dependent on the chosen pre-programme time period,  $s$ . In practice, we will need data from a period sufficiently prior to that of programme participation such that this future treatment is not anticipated (or does not affect the behaviour) in period  $s$ .

<sup>22</sup>The results are obtained using the probit specification. Similar results, available on request, came out using a logit specification. See also Heckman, Ichimura, Smith & Todd (1998).

<sup>23</sup>See appendix A for results on balancing properties for the different matching strategies.

<sup>24</sup>We also tried to **stratify** the common support of  $P(\mathbf{X})$  into a set of intervals. We took the mean difference in outcomes between the treated and non-treated observations within each interval and constructed an overall impact estimate as a weighted average of the interval impact estimates based on the fraction of the treated population in each interval. The results, available on request, are all in line with the ones presented below.

after training, this indicates the presence of a locking-in effect. The post-participation period is not sufficiently long to conclude whether the post-participation impact increases or decreases the unemployment rate of participants relative to non-participants.

**TABLE 5**  
 COMPARISON OF ESTIMATED CROSS-SECTIONAL TREATMENT EFFECTS  
 Dep. Variable: Unemployment Rate<sup>a</sup>  
 Full Sample, 305 Treated and Cross-Overs and 507 No-Shows and Real Controls

Month	Differences in Means	Experimental Estimates	Nearest Neighbour w/ Replacement <sup>b,c</sup>	Nearest Neighbour w/o Replacement <sup>b,d</sup>	Local Linear Matching <sup>e</sup>
July	<b>0.14</b> (0.03)	0.12 (0.10)	<b>0.10</b> (0.05)	<b>0.10</b> (0.04)	0.07 (0.04)
August	<b>0.12</b> (0.03)	0.08 (0.10)	0.04 (0.05)	0.06 (0.04)	0.03 (0.04)
September	<b>0.11</b> (0.03)	0.11 (0.10)	0.03 (0.05)	0.07 (0.04)	0.02 (0.04)
October	<b>0.10</b> (0.03)	0.06 (0.10)	0.01 (0.05)	0.07 (0.04)	0.04 (0.04)
November	<b>0.10</b> (0.03)	0.03 (0.10)	-0.01 (0.05)	0.06 (0.04)	0.05 (0.04)
December	0.06 (0.03)	0.04 (0.10)	-0.01 (0.05)	0.04 (0.04)	0.01 (0.04)
Average of July-Dec.	<b>0.10</b> (0.02)	0.07 (0.08)	0.03 (0.04)	<b>0.07</b> (0.03)	0.04 (0.04)

<sup>a</sup>No reweighting is done with censored observations.

<sup>b</sup>A caliper of 0.01 was used as a match tolerance. Std. errors. ignore estimation of the score.

<sup>c</sup>272 treated and 163 controls used.

<sup>d</sup>201 treated and controls used.

<sup>e</sup>Densities were estimated using a biweight kernel and a bandwidth proposed by Silverman (1986). The overlapping support region was determined using a 2 % trimming rule and a biweight kernel function. Matching done on log odds ratio. Std. errors are based on 1000 bootstraps with 100% resampling. See Heckman et al. (1997) for details. 277 treated and 503 controls used.

Generally speaking, the same picture emerges in table 6 where we focus on the D-in-D estimators. The impact estimates from the matching estimators are statistically significant (2 out of 3) in the month following participation, and then decline afterwards. The picture is thus still the same; there is evidence of a decline in the effect of training, once again supporting the hypothesis of a strong locking-in effect and a negligible post-programme effect.

**TABLE 6**

COMPARISON OF ESTIMATED DIFFERENCE IN DIFFERENCE TREATMENT EFFECTS  
 Dep. Variable: Unemployment Rate<sup>a</sup>

Full Sample, 305 Treated and Cross-Overs and 507 No-Shows and Real Controls

Month	Differences in Means	Nearest Neighbour w/ Replacement <sup>b,c</sup>	Nearest Neighbour w/o Replacement <sup>b,d</sup>	Local Linear Matching <sup>e</sup>
July	<b>0.13</b> -0.04	0.08 (0.05)	<b>0.09</b> (0.04)	<b>0.09</b> (0.04)
August	<b>0.12</b> (0.04)	0.02 (0.06)	0.06 (0.05)	0.06 (0.05)
September	<b>0.10</b> (0.04)	0.01 (0.06)	0.06 (0.05)	0.05 (0.05)
October	<b>0.09</b> (0.04)	-0.01 (0.06)	0.06 (0.05)	0.06 (0.05)
November	<b>0.09</b> (0.04)	-0.03 (0.06)	0.06 (0.05)	0.08 (0.05)
December	0.05 (0.04)	-0.03 (0.06)	0.04 (0.05)	0.04 (0.05)
Average of July-Dec.	<b>0.10</b> (0.03)	0.00 (0.05)	0.06 (0.04)	0.06 (0.04)

<sup>a</sup>No reweighting is done with censored observations.

<sup>b</sup>A caliper of 0.01 was used as a match tolerance. Std. errors. ignore estimation of the score.

<sup>c</sup>272 treated and 163 controls used.

<sup>d</sup>201 treated and controls used.

<sup>e</sup>Densities were estimated using a biweight kernel and a bandwidth proposed by Silverman (1986). The overlapping support region was determined using a 2 % trimming rule and a biweight kernel function. Matching done on log odds ratio. Std. errors are based on 1000 bootstraps with 100% resampling. See Heckman et al. (1997) for details. 277 treated and 503 controls used.

## 8 Conclusion and discussion

In this paper, we have formulated and estimated experimental and non-experimental estimators on experimental data in order to evaluate the effect of labour market training. The motivation for this approach is the presence of no-shows and cross-overs in the data; phenomena that occur in practically all social experiments, see e.g. Heckman et al. (1999). The impact parameters are estimated using data from a Danish social experiment conducted in 1994. The data consists of unemployed applicants for labour market training courses. These applicants were randomised into a treatment and a control group. Subsequently, 52% in the treatment group received treatment, but so did 22% in the control group.

Using a variety of estimation strategies, we find that the result obtained is very robust; training increases the subsequent unemployment rate of participants. However, we also find evidence that this is mostly due to a locking-in effect, since the effect disappears over time. However, the conclusion that training is bad for individual employment histories relative to no activity at all is premature. AMU training leave individual unemployment



relatively unchanged compared to the alternative, which might just indicate that other labour market policies are as efficient. However, most other studies on the Danish labour market tend to agree that most elements of the Danish active labour market policy actually increase the unemployment duration of participants, so we would be tempted to conclude that AMU training also increases individual unemployment. However, this is still not necessarily bad, if participants earn higher wages when they finally find jobs, but this issue is left for future research.

A number of lessons for future design of social experiments may be derived. First, it is important to obtain all relevant data for the experiment. Here, we particularly lack the information on the exact timing of the training programme, something which is crucial for the appropriate specification of the dynamic processes of training and job search. Had we had access to this information, we could have estimated a duration model, which would have enabled us to identify and estimate the locking-in and post programme effects directly, and with access to wage information we could even estimate the effect on subsequent wages and employment duration.

Secondly, the handling of compliance problems is important. Obviously, we would like to keep the control group 'pure', in the sense that non-participation should be enforced in the control group. However, no-shows are a fact of the training system which must be allowed for if a credible evaluation is to be performed. The fact that no-shows invalidate the basic experimental setup just means we should be more careful in making identifying assumptions and that we should check the robustness of our evaluation results.

Finally, the scale of the experiment should probably be slightly larger. In fact, in the Danish (Nordic) case, it would be highly useful to perform an 'reverse' experiment, where a large group of individuals is exposed to a treatment, which consists of exclusion from participation in active labour market programmes.

## References

- Barnow, B., Cain, G. & Goldberger, A. (1980), Issues in the analysis of selectivity bias, *in* E. Stromsdorfer & G. Farkas, eds, 'Evaluation Studies Review Annual', San Fransico: Sage, pp. 290–317.
- Björklund, A. (1993), The swedish experience, *in* K. Jensen & P. K. Madsen, eds, 'Measuring Labour Market Measures', Danish Ministry of Labour, pp. 243–263.
- Björklund, A. (1994), 'Evaluations of swedish labour market policy', *International Journal of Manpower* **15**, 16–31.
- Björklund, A. & Regnér, H. (1996), Experimental evaluation of european labour market policy, *in* Schmid, O'Reilly & Schumann, eds, 'International Handbook of Labour Market Policy and Evaluation', Edward Elgar, UK.
- Bonnal, L., Fougere, D. & Serandon, A. (1997), 'Evaluating the impact of french employment policies on individual labour market histories', *Review of Economic Studies* **64**, 683–713.
- Cochran, W. & Rubin, D. (1973), 'Controlling bias in observational studies', *Sankhya* **35**, 417–446.
- Dehejia, R. H. & Wahba, S. (1999), 'Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs', *Journal of the American Statistical Association* **94**(448), 1053–1062.
- Delander, L. (1978), 'Studier kring den arbetsformetlande verksamheten', *SOU* **60**.
- Harkman, A., Jansson, F. & Tamas, A. (1996), 'Effects, defects and prospects: An evaluation of labour market training in sweden', *Mimeo* .
- Heckman, J., Hohmann, N. & Smith, J. (2000), 'Substitution and dropout bias in social experiments: A study of an influential social experiment', *The Quarterly Journal of Economics* (May), 651–694.
- Heckman, J., Ichimura, H., Smith, J. & Todd, P. (1998), 'Characterizing selection bias using experimental data', *Econometrica* **66**(5), 1017–1098.
- Heckman, J., Ichimura, H. & Todd, P. (1997), 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *Review of Economic Studies* **64**, 605–654.
- Heckman, J., Ichimura, H. & Todd, P. (1998), 'Matching as an econometric evaluation estimator', *Review of Economic Studies* **65**, 261–294.
- Heckman, J. J. (1978), 'Dummy endogenous variables in a simultaneous equation system', *Econometrica* **46**, 931–959.
- Heckman, J. J. & Smith, J. A. (1999), 'The pre-program earnings dip and the determinants of participation in a social program: Implications for simple program evaluation strategies', *Economic Journal* **109**(457), 313–348.

- Heckman, J., LaLonde, R. & Smith, J. (1999), The economics and econometrics of active labor market programs, *in* A. Ashenfelter & D. Card, eds, 'Handbook of Labor Economics, Vol.', Vol. 3, Elsevier Science, Amsterdam, pp. 1865–2097.
- Jensen, P., Pedersen, P. J., Smith, N. & Westergaard-Nielsen, N. C. (1993), The effects of labour market training on wages and unemployment: Some danish results, *in* P. J. Bunzel, H. & N. Westergård-Nielsen, eds, 'Panel Data and Labour Market Dynamics', North Holland.
- Kraus, F., Puhani, P. & Steiner, V. (1997), 'Employment effects of publicly financed training programs - the east german experience', *Discussion Paper, ZEW* **97-33**.
- Langager, K. (1996), 'Ledige på kusus (unemployed in classroom training)', *Socialforsknings Instituttet* **96:9**.
- Raaum, O., Torp, H. & Goldstein, H. (1994), *Experiments in Manpower Policy Evaluation: The Case for Simple Estimators? Experience from a Norwegian Study of Labour Market Training*, Oslo.
- Richardsson, K. & Berg, G. J. V. D. (2002), The effect of vocational employment training on the individual transition rate from unemployment to work. IFAU Working Paper 2002:8.
- Rosenbaum, P. R. & Rubin, D. B. (1985), 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score', *The American Statistician* **39**(1), 33–38.
- Rosenbaum, P. & Rubin, D. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.
- Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall/CRC.
- Torp, H., Raaum, O., Hernæs, E. & Goldstein, H. (1993), The first norwegian experiment, *in* K. Jensen & P. K. Madsen, eds, 'Measuring Labour Market Measures', Danish Ministry of Labour, pp. 97–140.
- Van Den Berg, G. J. & der Klaauw, B. V. (2001), Counseling and monitoring of unemployed workers: Theory and evidence from a controlled social experiment. IZA Discussion Paper 374.
- Westergaard-Nielsen, N. C. (1993), Effects of training: A fixed-effect model, *in* P. K. M. K. Jensen, ed., 'Measuring Labour Market Measures', Vol. 93-3, Danish Ministry of Labour.
- White, M. & Lakey, J. (1992), 'The restart effect: Does active labour market policy reduce unemployment?', *Policy Studies Institute, London* .

## A Balancing Score

Our use of the propensity score in reducing the curse of dimensionality should balance the distribution of covariates,  $\mathbf{X}$ , see Rosenbaum & Rubin (1983). This property is investigated informally in the present section. We do not investigate the full distribution of covariates, but focus merely on the first two moments (see also Dehejia & Wahba (1999)). Table A presents the characteristics of the matching estimators implemented. We note that a large subset of the covariates did already balance before any reweighting was done, see table 2; a property that survives the matching. The standardized differences are calculated as (see Rosenbaum & Rubin (1985))

$$100 * \frac{(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_{0M})}{\sqrt{(S_1^2 + S_{0M}^2)/2}}$$

The tables below demonstrate the balancing properties. Generally, the Nearest Neighbour with replacement algorithm does the poorest job in balancing the covariates with 6 out of the 40 having a bias (numerically) larger than 10%; two of them capturing important labour market dynamics and one the important information of being recommended to participate (table A1). The Nearest Neighbour without replacement does a slightly better job in balancing the covariates: 5 out of the 40 have an estimated bias larger than 10% but only one of the important labour market variables are among the 5 (table A2). The least imbalance is produced by the kernel based method (table A3) where non of the estimated biases are above 10%. All in all we conclude that the kernel based strategy produces a group of matched controls with characteristics near identical to that of the treated. This is achieved with just a simple linear index of the covariates ignoring any cross or higher order terms. Notice further, that the nearest neighbour matching without replacement achieves this rather poor balancing property at the cost of discarding nearly 40 % of the treated, and that for both of the two types of nearest neighbour more cross terms and higher order terms properly would be needed to be added in the balancing score in order for these to produce as balanced matched samples as the kernel based.

**TABLE A1**  
**COVARIATE IMBALANCE IN NEAREST NEIGHBOUR W/ REPLACEMENT SAMPLE**  
Means and Standardized Differences in Percentage Points<sup>a</sup>

Variables	Treated Means	Matched Means	% Bias
Aabenraa	0.15	0.17	-3.00
Aalborg	0.22	0.23	-0.88
Odense	0.21	0.19	3.68
Randers	0.26	0.25	2.53
Male	0.76	0.78	-3.48
Woman with 'R=0'- status	0.10	0.08	3.85
Immigrant	0.06	0.05	6.43
Age below 25	0.30	0.29	1.61
Age 25 to 29	0.24	0.24	-1.72
Age 30 to 39	0.25	0.28	-7.47
Age 40 to 49	0.16	0.16	-1.00
Age 50 to 54	0.04	0.02	14.80
Fewer than 9 years of schooling	0.11	0.11	0.00
High School education	0.05	0.05	0.00
Vocational education	0.36	0.39	-6.07
Short further education	0.02	0.01	5.47
College education and beyond	0.03	0.01	7.82
Living with a partner	0.64	0.61	4.55
Partner unemployed during last year	0.27	0.31	-8.88
Child age less than 6 present in household	0.24	0.24	0.00
Self-reported labour market experience. years	10.85	10.15	7.40
One job spell in 1993	0.59	0.65	-11.36
Two job spells in 1993	0.12	0.13	-1.12
More than two job spells in 1993	0.04	0.04	-3.73
One job spell in 1994 prior to RA	0.37	0.46	-17.94
More than one job in 1994 prior to RA	0.04	0.03	8.21
Last job in public sector	0.06	0.07	-4.61
Last job in private sector	0.40	0.38	4.50
Self-reported health status. medium	0.10	0.08	5.08
Self-reported health status. poor	0.02	0.01	12.22
Applied based on expected employment gain	0.75	0.75	0.85
Encouraged to participate	0.15	0.19	-11.67
Forced to participate	0.02	0.02	-2.61
Unemployment rate in month prior to RA	0.59	0.57	4.20
No Unemployment Insurance	0.12	0.17	-15.67
UI in Metal	0.08	0.10	-5.08
UI in KAD (female)	0.04	0.05	-3.44
UI in Construction	0.02	0.03	-2.40
UI other place	0.29	0.27	5.72
R	0.72	0.71	0.81

<sup>a</sup>Sample size is 305 treated and 272 controls

**TABLE A2**  
**COVARIATE IMBALANCE IN NEAREST NEIGHBOUR W/O REPLACEMENT SAMPLE**  
Means and Standardized Differences in Percentage Points<sup>a</sup>

Variables	Treated Means	Matched Means	% Bias
Aabenraa	0.12	0.19	-19.08
Aalborg	0.20	0.23	-7.27
Odense	0.20	0.18	5.07
Randers	0.27	0.25	4.53
Male	0.73	0.73	0.00
Woman with 'R=0'- status	0.12	0.11	1.54
Immigrant	0.05	0.05	2.23
Age below 25	0.32	0.27	10.91
Age 25 to 29	0.20	0.24	-8.37
Age 30 to 39	0.26	0.27	-1.12
Age 40 to 49	0.16	0.17	-1.33
Age 50 to 54	0.03	0.04	-2.62
Fewer than 9 years of schooling	0.13	0.11	4.55
High School education	0.04	0.04	0.00
Vocational education	0.32	0.40	-17.64
Short further education	0.03	0.02	6.38
College education and beyond	0.02	0.02	0.00
Living with a partner	0.66	0.64	3.12
Partner unemployed during last year	0.25	0.26	-3.41
Child age less than 6 present in household	0.27	0.23	8.02
Self-reported labour market experience. years	10.70	11.06	-3.54
One job spell in 1993	0.55	0.60	-10.06
Two job spells in 1993	0.12	0.12	0.00
More than two job spells in 1993	0.04	0.02	8.42
One job spell in 1994 prior to RA	0.41	0.40	2.02
More than one job in 1994 prior to RA	0.04	0.04	-2.47
Last job in public sector	0.05	0.05	0.00
Last job in private sector	0.36	0.39	-6.14
Self-reported health status. medium	0.10	0.09	1.68
Self-reported health status. poor	0.02	0.01	7.11
Applied based on expected employment gain	0.74	0.76	-3.43
Encouraged to participate	0.12	0.15	-8.61
Forced to participate	0.03	0.02	6.38
Unemployment rate in month prior to RA	0.57	0.57	0.60
No Unemployment Insurance	0.14	0.12	4.40
UI in Metal	0.08	0.09	-3.57
UI in KAD (female)	0.04	0.08	-14.42
UI in Construction	0.02	0.03	-5.83
UI other place	0.29	0.26	4.45
R	0.64	0.65	-2.08

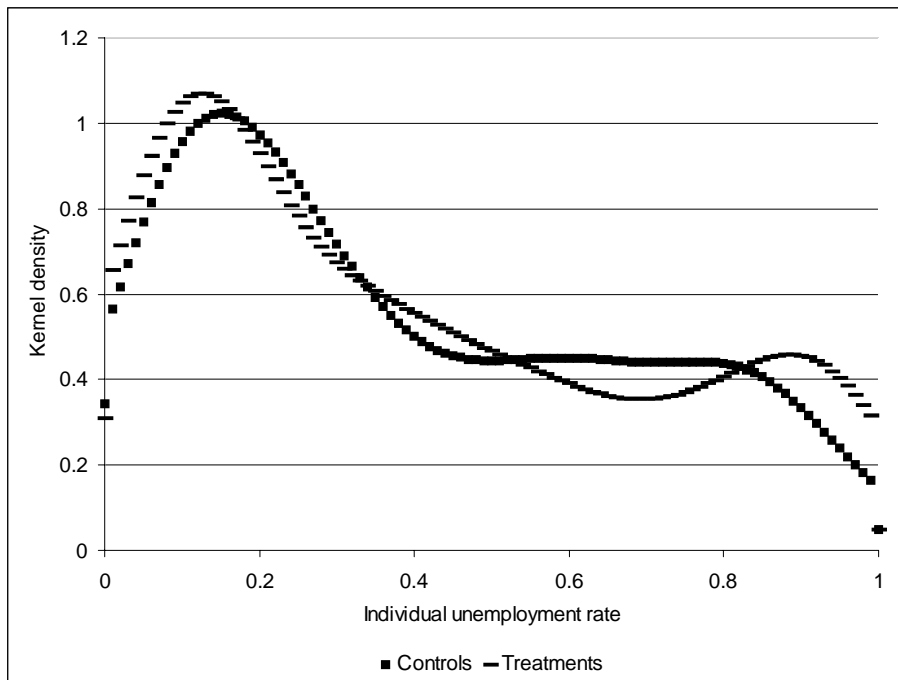
<sup>a</sup>Sample size is 201 treated and controls

**TABLE A3**  
**COVARIATE IMBALANCE IN LOCAL LINEAR MATCHING SAMPLE**  
Means and Standardized Differences in Percentage Points<sup>a</sup>

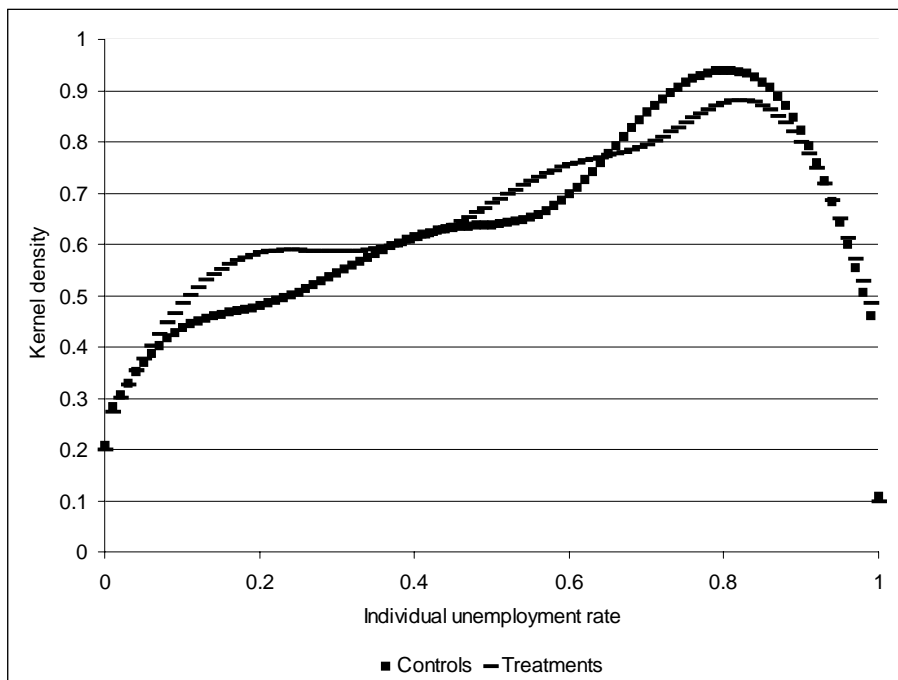
Variables	Treated Means	Matched Means	% Bias
Aabenraa	0.16	0.17	-4.76
Aalborg	0.27	0.23	0.09
Odense	0.19	0.18	3.38
Randers	0.26	0.26	-1.19
Male	0.75	0.76	-1.57
Woman with 'R=0'- status	0.10	0.09	5.22
Immigrant	0.06	0.07	-3.53
Age below 25	0.30	0.30	-2.03
Age 25 to 29	0.24	0.21	6.47
Age 30 to 39	0.25	0.28	-7.11
Age 40 to 49	0.15	0.15	0.79
Age 50 to 54	0.04	0.03	4.69
Fewer than 9 years of schooling	0.12	0.13	-3.97
High School education	0.05	0.03	7.66
Vocational education	0.35	0.35	0.61
Short further education	0.02	0.02	-0.50
College education and beyond	0.03	0.02	0.63
Living with a partner	0.62	0.64	-2.22
Partner unemployed during last year	0.26	0.26	-0.76
Child age less than 6 present in household	0.22	0.22	1.02
Self-reported labour market experience. years	10.73	10.56	1.71
One job spell in 1993	0.57	0.58	-2.64
Two job spells in 1993	0.14	0.13	1.71
More than two job spells in 1993	0.04	0.03	5.16
One job spell in 1994 prior to RA	0.38	0.37	1.27
More than one job in 1994 prior to RA	0.04	0.04	-1.22
Last job in public sector	0.04	0.06	-8.17
Last job in private sector	0.41	0.39	5.28
Self-reported health status. medium	0.10	0.10	-0.99
Self-reported health status. poor	0.02	0.02	1.95
Applied based on expected employment gain	0.74	0.72	4.54
Encouraged to participate	0.14	0.13	2.42
Forced to participate	0.03	0.04	-6.60
Unemployment rate in month prior to RA	0.58	0.60	-5.43
No Unemployment Insurance	0.12	0.13	-5.81
UI in Metal	0.09	0.07	5.75
UI in KAD (female)	0.04	0.04	0.00
UI in Construction	0.02	0.02	-1.31
UI other place	0.30	0.28	4.40
R	0.69	0.70	-1.49

<sup>a</sup>Sample size is 207 treated and 503 controls

## B Kernel Densities



Kernel density estimates of individual unemployment rates, post treatment period.



Kernel density estimates of individual unemployment rates, pre-treatment period.



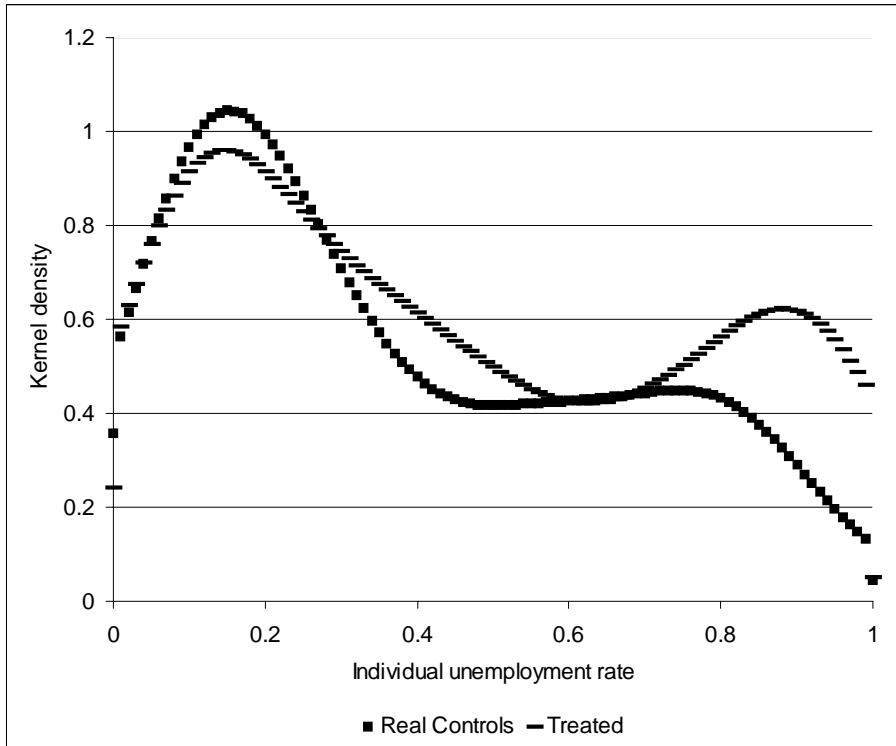


Figure 2: Kernel density estimates of individual unemployment rates, post-treatment period.

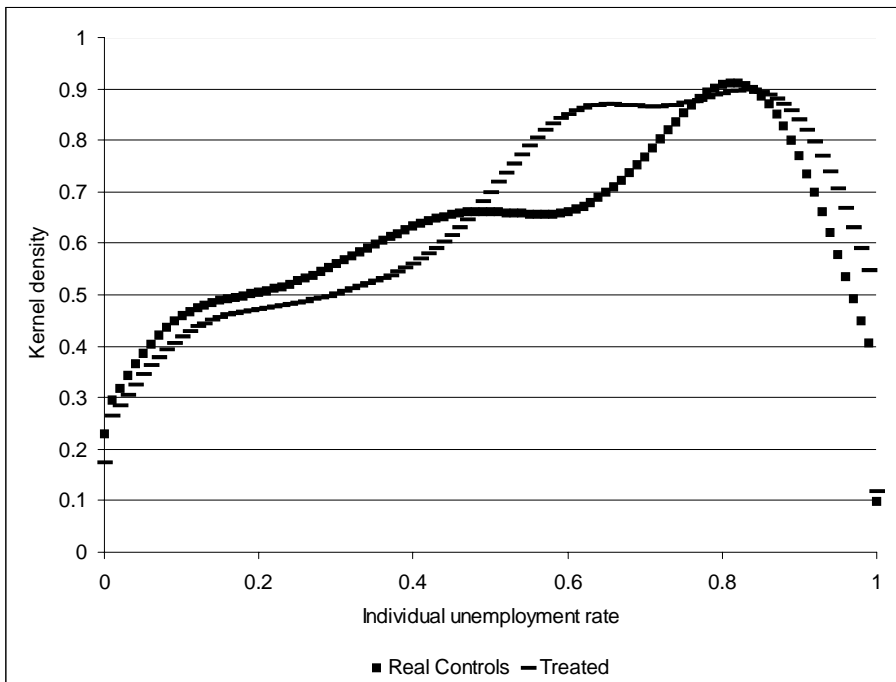


Figure 3: Kernel density estimates of individual unemployment rates, pre-treatment period.

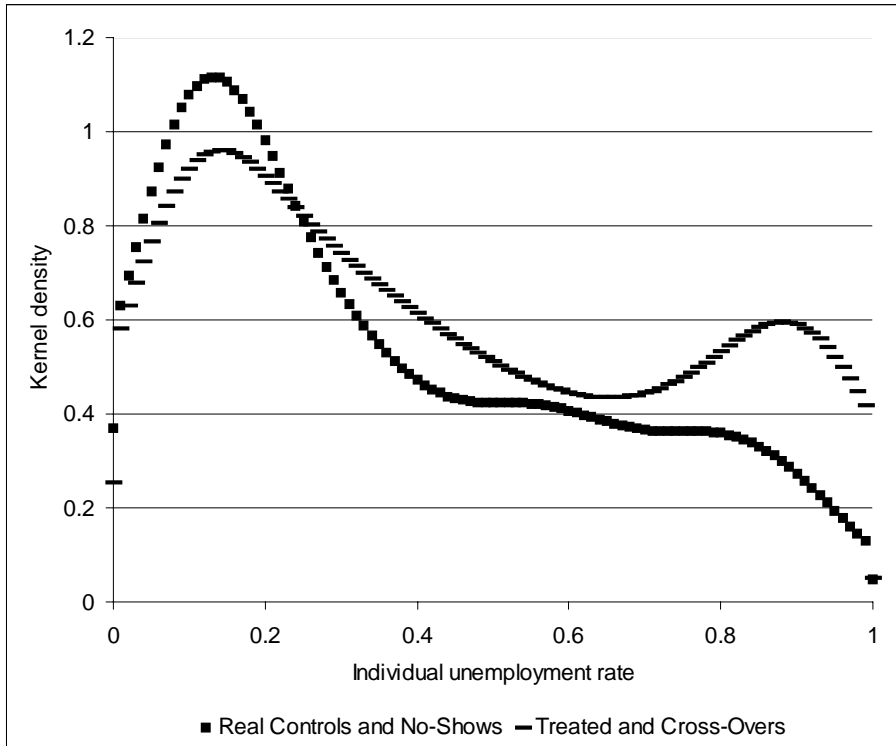


Figure 4: Kernel density estimates of individual unemployment rates, post-treatment period.

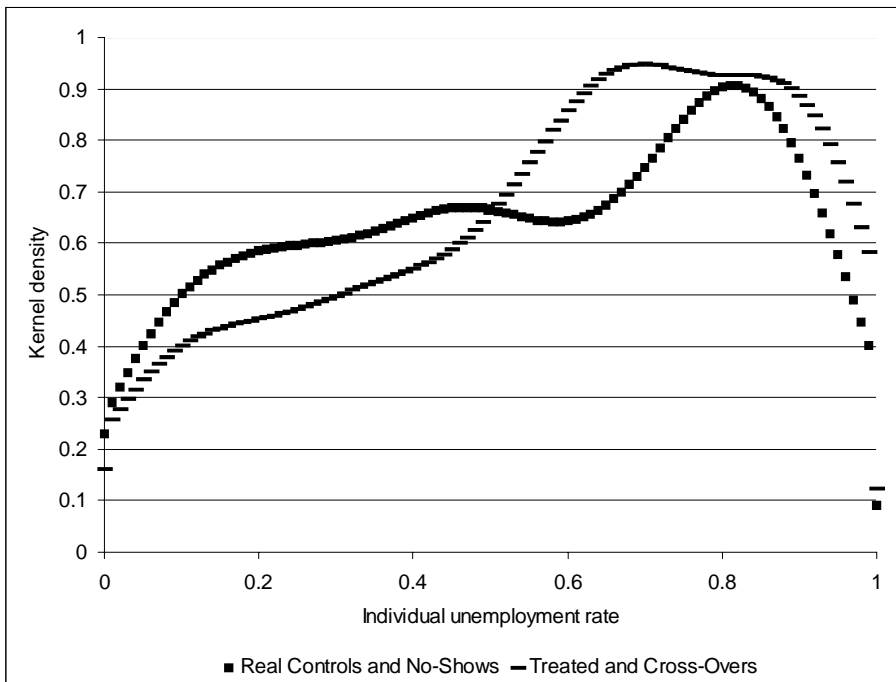


Figure 5: Kernel density estimates of individual unemployment rates, pre-treatment period.

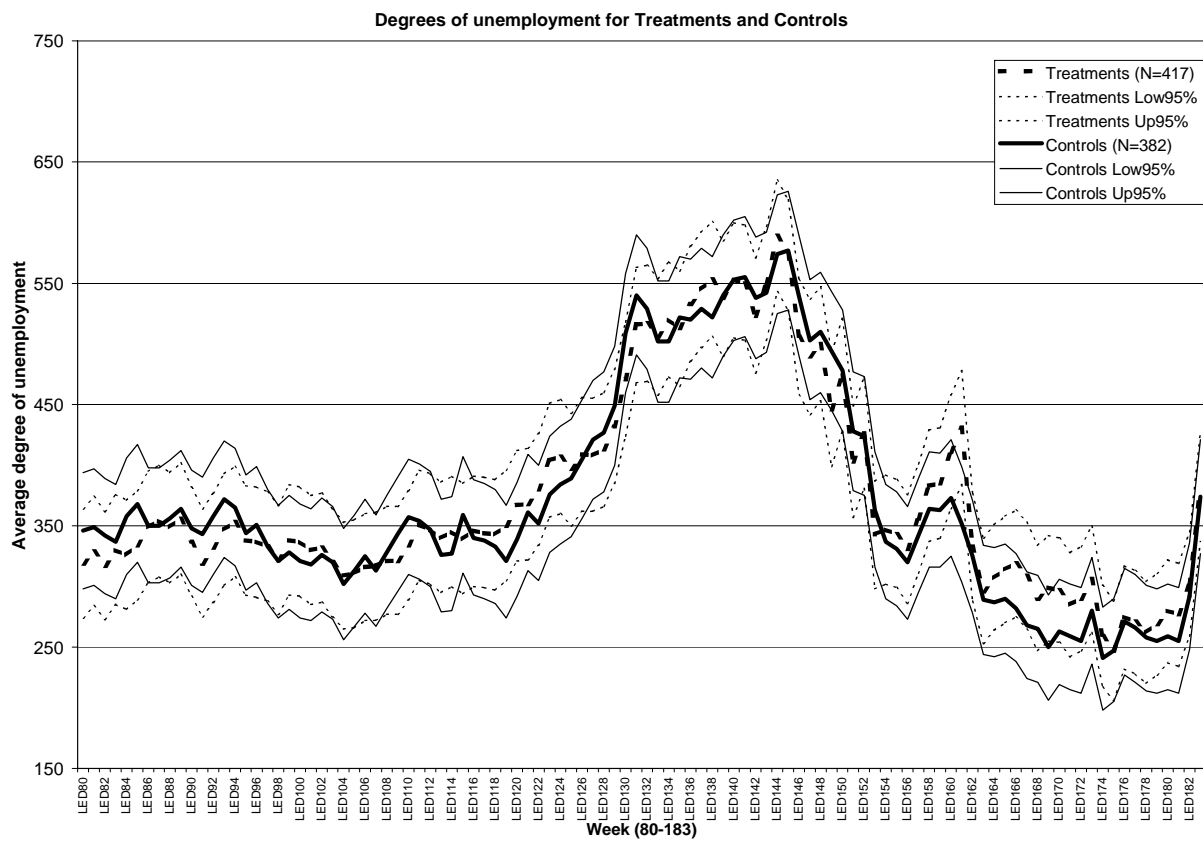


Figure 6: Mean individual weekly unemployment rates.

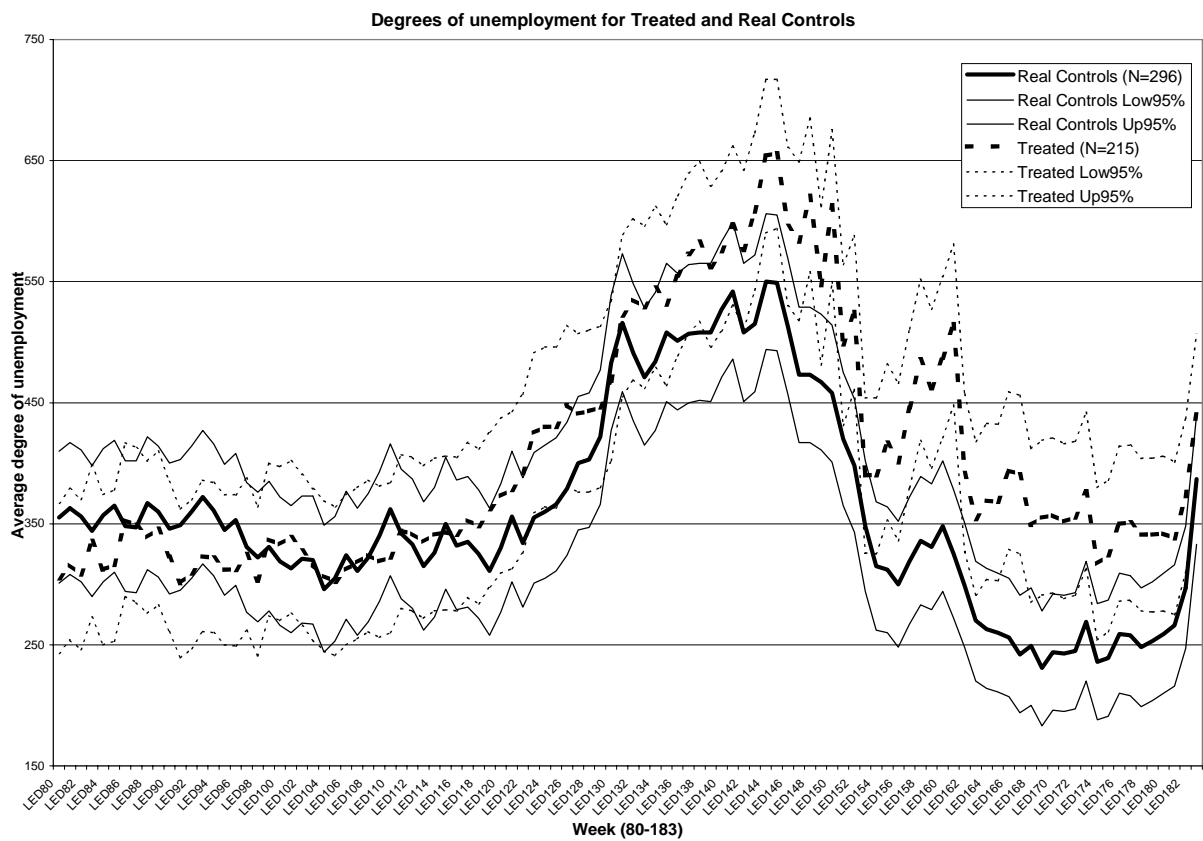


Figure 7: Mean individual weekly unemployment rates.

## IZA Discussion Papers

No.	Author(s)	Title	Area	Date
701	L. Goerke J. B. Madsen	Earnings-Related Unemployment Benefits in a Unionised Economy	3	01/03
702	M. Ayhan Kose E. S. Prasad M. E. Terrones	How Does Globalization Affect the Synchronization of Business Cycles?	2	01/03
703	W. C. Horrace R. L. Oaxaca	New Wine in Old Bottles: A Sequential Estimation Technique for the LPM	6	01/03
704	M. Ransom R. L. Oaxaca	Intrafirm Mobility and Sex Differences in Pay	1	01/03
705	G. Brunello D. Checchi	School Quality and Family Background in Italy	2	01/03
706	S. Girma H. Görg	Blessing or Curse? Domestic Plants' Survival and Employment Prospects after Foreign Acquisitions	1	01/03
707	C. Schnabel J. Wagner	Trade Union Membership in Eastern and Western Germany: Convergence or Divergence?	3	01/03
708	C. Schnabel J. Wagner	Determinants of Trade Union Membership in Western Germany: Evidence from Micro Data, 1980-2000	3	01/03
709	L. Danziger S. Neuman	Delays in Renewal of Labor Contracts: Theory and Evidence	1	02/03
710	Z. Eckstein Y. Weiss	On the Wage Growth of Immigrants: Israel, 1990-2000	2	02/03
711	C. Ruhm	Healthy Living in Hard Times	3	02/03
712	E. Fehr J. Henrich	Is Strong Reciprocity a Maladaptation? On the Evolutionary Foundations of Human Altruism	5	02/03
713	I. Gang J. Landon-Lane M. S. Yun	Does the Glass Ceiling Exist? A Cross-National Perspective on Gender Income Mobility	2	02/03
714	M. Fertig	Educational Production, Endogenous Peer Group Formation and Class Composition – Evidence From the PISA 2000 Study	6	02/03
715	E. Fehr U. Fischbacher B. von Rosenblatt J. Schupp G. G. Wagner	A Nation-Wide Laboratory Examining Trust and Trustworthiness by Integrating Behavioral Experiments into Representative Surveys	7	02/03
716	M. Rosholm L. Skipper	Is Labour Market Training a Curse for the Unemployed? Evidence from a Social Experiment	6	02/03

An updated list of IZA Discussion Papers is available on the center's homepage [www.iza.org](http://www.iza.org).