The matching method for treatment evaluation with selective participation and ineligibles

Monica Costa Dias
Hidehiko Ichimura
Gerard van den Berg

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP35/10

The Matching Method for Treatment Evaluation with Selective Participation and Ineligibles

Monica Costa Dias*

Hidehiko Ichimura[†]

Gerard J. van den Berg[‡]

November 2010

Abstract

The matching method for treatment evaluation does not balance selective unobserved differences between treated and non-treated. We derive a simple correction term if there is an instrument that shifts the treatment probability to

*Cef.up, Faculty of Economics at the University of Porto; IFS and IZA. Address: IFS, 7 Ridgmount Street, London WC1E 7AE, UK. monica_d@ifs.org.uk

[†]University of Tokyo.

[‡]Alexander von Humboldt Professor of Econometrics and Empirical Economics, University of Mannheim; VU University Amsterdam, IFAU-Uppsala, CEPR, IZA, and IFS.

Keywords: propensity score, regression discontinuity, instrumental variable, policy evaluation, treatment effect, selection, job search assistance, subsidized work, youth unemployment.

Acknowledgements: We thank Richard Blundell, Xavier de Luna, Barbara Sianesi, and Petra Todd, for useful comments. We also thank Louise Kennerberg and Barbara Sianesi for help with the Swedish data. We gratefully acknowledge financial support from the ESRC and from IFAU-Uppsala. Costa Dias is thankful to Fundacao para a Ciencia e Tecnologia and the European Social Fund for financial support.

zero in specific cases. Within the same framework we also suggest a new test of the conditional independence assumption justifying matching. Policies with eligibility restrictions, where treatment is impossible if some variable exceeds a certain value, provide a natural application. In an empirical analysis, we exploit the age eligibility restriction in the Swedish Youth Practice subsidized work program for young unemployed, where compliance is imperfect among the young. Adjusting the matching estimator for selectivity changes the results towards making of subsidized work detrimental in moving individuals into employment.

1 Introduction

The matching method for treatment evaluation compares outcomes of treated and non-treated subjects, conditioning on observed individual and environment characteristics. Basically, the average treatment effect on the treated (ATT) is estimated by averaging observed outcome differences over the treated. The main assumption is that the conditioning ensures that the assigned treatment status is conditionally mean independent from the potential outcomes (this is usually known as "the Conditional Independence Assumption" or, in short, CIA, although in fact it respects to mean independence).[1]

The method is intuitive, as it mimics randomized experiments: the distributions of behavioral determinants and indicators are balanced as closely as possible over treated and non-treated, using observational data. The use of the method has improved the policy evaluation practice by clarifying the importance of common support restrictions for the distribution of conditioning variables. By now, it is a common tool for the analysis of active labor market policies (ALMP) and programs (see e.g. the survey

[1] See e.g. Cochrane and Rubin (1973), Rosenbaum and Rubin (1983), and Heckman, Ichimura, and Todd (1998).

in Kluve, 2006). However, matching has the well-recognized limitation that it does not ensure the balancing of the distribution of unobservable determinants of both treatment assignment and outcomes among treated and non-treated. When incapable to balance unobservables, matching may produce biased estimates of the treatment effects.

The first contribution of this paper deals with this problem by developing an estimation method for the average treatment on the treated robust to violations in the conditional independence assumption justifying matching. The idea is to correct the matching estimate with a measure of the bias due to selection on unobservables. Key to the estimation of such correction term is the availability of an instrument capable of driving participation to zero at certain of its (possibly limiting) values while keeping the selection mechanism partly unexplained at other parts of its distribution. Like the matching methods, our approach matches the distribution of observed variables between treated and non-treated groups, thus effectively combining matching with the exogenous variation provided by an instrument to balance unobservables.

Alternative approaches in order to correct matching estimators for selection problems typically assume that the relevant unobserved variables have additive effects on the potential outcomes (see Heckman and Robb, 1985, and Andrews and Schafgans, 1998). The popular conditional difference-in-differences estimator (Heckman, Ichimura, Smith, and Todd, 1998) is also based on this. By contrast, our approach does not require additivity.

Within the same framework we also suggest a new test of the CIA. In the presence of a valid instrument, satisfying the condition introduced above, the CIA holds if and only if the correction term is zero. Thus, testing the validity of the CIA is equivalent to test the statistical significance of the correction term.

The second contribution of this paper is to show that there are important empirical applications for this method. Consider, for instance, the case of treatment evaluation in the context of ALMP for unemployed workers. It has been recognized that individual characteristics and employment history may not capture the full range of skills and motivation that explain both treatment participation and employment-related outcomes.[2] However, many programs use clear eligibility rules based on observed variables and often including boundary restrictions. Often, such rules can be exploited to construct an instrument capable of moving subjects in and out of treatment while otherwise being unrelated to the potential outcome(s) of interest. The ideal setting for the application of our method is created in the presence of boundary restrictions on personal characteristics and conditions such that there is full non-participation at certain values of the instrument while compliance is imperfect at other values.[3]

This is a relevant setting. It is a common feature of ALMP to restrict eligibility to individuals aged above or below a certain age, or to individuals with a certain minimum or maximum amount of education, and/or to individuals with a certain minimum amount of labor market experience (see e.g. Kluve, 2006). If imperfect compliance among the eligible individuals is selective then the matching approach cannot be used. We propose overcoming this limitation by exploiting the eligibility boundary restriction within the matching framework.

Our approach is related to Battistin and Rettore (2008), who consider a specific partially fuzzy discontinuity design where eligibility rules preclude participation on one side of a threshold for a certain variable and allow - but do not impose - it on

[2]For example, Card and Sullivan (1988), Gritz (1993), Bonnal, Fougère and Sérandon (1997) and Richardson and Van den Berg (2001) argue that this can be expected to play a major role in the empirical evaluation of ALMP, and their estimation results confirm this. Van den Berg, Van der Klaauw and Van Ours (2004) contain similar findings for the effect of punitive sanctions for welfare recipients.

[3]Here, the word "compliance" is used in a statistical sense, meaning that some of the individuals who, according to the policy design, are eligible for treatment end up in the non-treated subpopulation.

the opposite side (so non-compliance affects outcomes only on one side of the discontinuity). To identify a Local Average Treatment Effect (LATE), they only need the continuity assumption that is characteristic of sharp regression discontinuity (RD) designs. They notice that the identified parameter is a LATE and they derive the similarities with the "Bloom setting": a fully experimental setting with non-compliance on the treated side only.[4]

We derive a similar estimator, but our derivations are from a matching perspective, and, accordingly, our quantity of interest is the average treatment effect on the treated (ATT). The underlying assumptions are not identical and, as a result, the applicability also differs. Our assumption of ignorability of an instrumental variable Z conditional on observables X is a typical exclusion restriction. It implies Battistin and Rettore (2008)'s continuity assumption around the threshold point. At the same time, we avoid the discontinuity in participation that they use for identification. Our estimator can be seen as a discretized version of their RD estimator. Both estimators lead to a test of the CIA. While deriving our proposed testing procedure (Subsection 2.2) we point out in what sense it differs from Battistin and Rettore (2008)'s approach. Our empirical application (Section 3) illustrates how our approach can be applied in a case where the RD approach is not appropriate.

We empirically assess our approach by evaluating a major Swedish program aimed at helping unemployed individuals aged below 25 to find work, the Youth Practice (YP).[5] YP is a subsidized work program designed for short-term unemployed individ-

[4]The idea of exploiting one-sided compliance to deal with selective participation has some history in the analysis of treatment effects on duration outcomes in Mixed Proportional Hazard types of models with endogenous treatments. See Bijwaard and Ridder (2005) and Abbring and Van den Berg (2005).

[5]There is an increasing awareness that youth unemployment may be a serious problem for society despite the fact that youth unemployment durations are relatively short. This is because of the prevalence of psychological and labor-market scarring effects which may have long-run implications

uals aged below 25. The program is not compulsory, being one among a number of alternative treatments. This means that compliance is imperfect on the lower side of the age-eligibility threshold. We may therefore apply our selectivity-adjusted matching estimator using age as the instrument. The subpopulation of non-treated includes those below 25 who do not participate as well as those 25 and above. Participation is not sharply discontinuous at age 25 but declines gradually before age 25. This is not a problem for our method but could complicate the application of regression-discontinuity methods. The non-compulsory nature of the program among eligibles may raise difficulties for matching to balance unobservables. We use our correction factor to assess whether this is in fact the case and to eliminate the potential selection bias.

The Swedish YP has been evaluated before (see e.g. White and Knight, 2002, Larsson, 2003, Forslund and Nordström Skans, 2006, for results). It is of particular interest that existing YP evaluations are based on the matching approach. We find that adjusting the matching estimator for selectivity changes the results to become negative when the outcome of interest is outflow into employment.

In Section 2 we develop a formal framework for the analysis. We define the objects of interest and we derive the selectivity-adjusted matching estimator. In Section 3 we discuss the Swedish YP program, estimation details, data and estimates. Section 4 concludes.

for the productivity of those affected (see e.g. Burgess et al., 2003).

2 A correction term to matching

2.1 Identification of the ATT in case of selective participation and ineligibles

In what follows, we adopt standard counterfactual notation where Y_0 and Y_1 are individual potential outcomes associated with being assigned to non-treatment and treatment, respectively. The binary indicator, D, denotes the actual treatment status, where we use the terms "participation" and "treatment" to denote $D = 1$ and "non-participation" and "control" to denote $D = 0$. The vector X contains conditioning variables. The actual outcome Y satisfies $Y = DY_1 + (1 - D)Y_0$.

We are interested in the Average Treatment Effect on the Treated (ATT):

$$\text{ATT} = \text{E}[Y_1 - Y_0 \mid D = 1].$$

Clearly,

$$\begin{aligned} \text{ATT} &= \text{E}_{X|D=1}\text{E}[Y_1 - Y_0 \mid X, D = 1] = \\ &\quad \text{E}_{X|D=1}\text{E}[Y_1 \mid X, D = 1] - \text{E}_{X|D=1}\text{E}[Y_0 \mid X, D = 1] \end{aligned} \tag{1}$$

where the expectations $\text{E}_{X|D=1}$ are taken over the distribution of X among the treated. Under the unconfoundedness assumption or Conditional Independence Assumption (CIA) stating that $Y_0 \perp D \mid X$, the ATT is identified and can be estimated using a matching method. We do *not* make such an assumption, because we do not rule out that for given X, the actual treatment assignment at the individual level is related to the potential gain of the treatment (we refer to the latter possibility as "selection on unobservables"). Instead, we assume that there exists a variable Z with the following two features,

1. $Y_0 \perp Z \mid X$;

2. There exists a set of points $\{z^*, z^{**}\}$ in the domain of Z where

$$\mathrm{P}\left[D = 1 \mid X, Z = z^*\right] = 0 \quad \text{and} \quad 0 < \mathrm{P}\left[D = 0 \mid X, Z = z^{**}\right] < 1$$

for all X.

Assumption 1 states that Z does not explain Y_0 when conditioning on the explanatory variables, X. Assumption 2 states that D is a non-trivial function of Z after conditioning on X. More specifically, the variation of D with Z satisfies two properties: first, participation can be driven to zero at certain parts of the distribution of Z, and second, participation is not deterministic over other parts of the distribution. If the participation probability is zero then we call the individual ineligible. Assumptions 1 and 2 do not rule out that participation is selective. In particular, if $Z = z^{**}$, then D may depend on Y_0 even if we condition on X.[6] Assumptions 1 and 2 can be called an exclusion restriction and an "informative instrument" assumption, so it is natural to call Z an instrumental variable. Notice that Assumption 2 can be verified empirically, whereas Assumption 1 requires an external justification.

In the above expression for ATT, the term $\mathrm{E}_{X|D=1}\mathrm{E}[Y_1 \mid X, D = 1]$ is directly identified from the mean observed outcome among the treated. The challenge is to identify the mean counterfactual outcome $\mathrm{E}_{X|D=1}\mathrm{E}[Y_0 \mid X, D = 1]$.

Under Assumption 1,

$$\begin{aligned} \mathrm{E}\left[Y_0 \mid X\right] &= \mathrm{E}\left[Y_0 \mid X, Z\right] \\ &= \mathrm{E}\left[Y_0 \mid X, Z, D = 0\right] \mathrm{P}\left[D = 0 \mid X, Z\right] + \\ &\quad\ \mathrm{E}\left[Y_0 \mid X, Z, D = 1\right] \mathrm{P}\left[D = 1 \mid X, Z\right]. \end{aligned} \tag{2}$$

Since this relationship holds for all possible values of Z, and in particular for $Z = z^*$, Assumption 2 ensures that

$$\mathrm{E}\left[Y_0 \mid X\right] = \mathrm{E}\left[Y_0 \mid X, Z = z^*, D = 0\right]. \tag{3}$$

[6]We tacitly make other standard assumptions like SUTVA and common support.

On the other hand, the following decomposition always yields,

$$\begin{aligned} \mathrm{E}\left[Y_0 \mid X\right] &= \mathrm{E}\left[Y_0 \mid X, D=0\right] \mathrm{P}\left[D=0 \mid X\right] + \\ &\quad \mathrm{E}\left[Y_0 \mid X, D=1\right] \mathrm{P}\left[D=1 \mid X\right] \end{aligned}$$

implying

$$\begin{aligned} &\mathrm{E}\left[Y_0 \mid X, D=1\right] \\ &\quad = \frac{\mathrm{E}\left[Y_0 \mid X\right] - \mathrm{E}\left[Y_0 \mid X, D=0\right] \mathrm{P}\left[D=0 \mid X\right]}{\mathrm{P}\left[D=1 \mid X\right]} \\ &\quad = \frac{\mathrm{E}\left[Y_0 \mid X, Z=z^*, D=0\right] - \mathrm{E}\left[Y_0 \mid X, D=0\right] \mathrm{P}\left[D=0 \mid X\right]}{\mathrm{P}\left[D=1 \mid X\right]} \\ &\quad = \mathrm{E}\left[Y_0 \mid X, D=0\right] + \frac{\mathrm{E}\left[Y_0 \mid X, Z=z^*, D=0\right] - \mathrm{E}\left[Y_0 \mid X, D=0\right]}{1 - \mathrm{P}\left[D=0 \mid X\right]}. \end{aligned} \tag{4}$$

Equation (4) is an expression for the mean counterfactual outcome $\mathrm{E}[Y_0 \mid X, D=1]$ given X. The mean counterfactual outcome given X that is used in standard matching estimation, $\mathrm{E}\left[Y_0 \mid X, D=0\right]$, is corrected for individual selection on unobservables by the second term in line four of the equation.

The terms $\mathrm{E}\left[Y_0 \mid X, D=0\right]$ and $\mathrm{E}[Y_0 \mid X, Z=z^*, D=0]$ in the right-hand side of equation (4) are identified from the mean observed outcome among the controls at given X and the mean observed outcome among the ineligible controls at given X, respectively. Taken together, this implies that the mean counterfactual outcome $\mathrm{E}[Y_0 \mid X, D=1]$ given X is identified from equation (4). In turn, the mean counterfactual outcome $\mathrm{E}_{X|D=1}\mathrm{E}[Y_0 \mid X, D=1]$ unconditional on X is identified by averaging over the observable distribution of X given $D=1$. Hence, the ATT is identified. Notice that identification does not require any additivity assumption on the relationships between outcome, treatment, and instrument. Also, identification does not require the instrument to be discrete or to be continuous. In the next subsection we discuss in some detail how we may implement the estimator suggested by the above constructive identification proof.

An alternative but similar approach to identification and inference is based on the fact that in the absence of selection on unobservables, we can discard the ineligibles and instead use only the eligible controls to obtain the mean counterfactual outcome for the treated. In general, we can express the ATT as

$$\text{ATT} = \mathrm{E}_{[X,Z=z^{**}|D=1]}\,\mathrm{E}[Y_1 - Y_0 \mid X, Z = z^{**}, D = 1]$$

where z^{**} stands for all possible values of z^{**} satisfying Assumption 2. This expression for the ATT follows from the fact that $D = 1$ automatically implies that $Z = z^{**}$ for some z^{**} satisfying Assumption 2. We will now follow the above derivation of the identification of $\mathrm{E}[Y_0 \mid X, D = 1]$, where we now condition on $Z = z^{**}$ as well.

The mean no-treatment outcome at a specific point $(X, Z = z^{**})$ with a non-zero probability of treatment is

$$\begin{aligned}\mathrm{E}\left[Y_0 \mid X, Z = z^{**}\right] &= \mathrm{E}\left[Y_0 \mid X, Z = z^{**}, D = 0\right]\mathrm{P}\left[D = 0 \mid X, Z = z^{**}\right] \\ &\quad + \mathrm{E}\left[Y_0 \mid X, Z = z^{**}, D = 1\right]\mathrm{P}\left[D = 1 \mid X, Z = z^{**}\right]\end{aligned}$$

while Assumptions 1 and 2 ensure that

$$\begin{aligned}\mathrm{E}\left[Y_0 \mid X, Z = z^{**}\right] &= \mathrm{E}\left[Y_0 \mid X, Z = z^{*}\right] \\ &= \mathrm{E}\left[Y_0 \mid X, Z = z^{*}, D = 0\right].\end{aligned}$$

But then, the counterpart of (4) if conditioning on Z is

$$\begin{aligned}&\mathrm{E}\left[Y_0 \mid X, Z = z^{**}, D = 1\right] \\ &\quad = \mathrm{E}\left[Y_0 \mid X, Z = z^{**}, D = 0\right] \\ &\qquad + \frac{\mathrm{E}\left[Y_0 \mid X, Z = z^{*}, D = 0\right] - \mathrm{E}\left[Y_0 \mid X, Z = z^{**}, D = 0\right]}{\mathrm{P}\left[D = 1 \mid X, Z = z^{**}\right]}.\end{aligned} \tag{5}$$

The terms $\mathrm{E}\left[Y_0 \mid X, Z = z^{**}, D = 0\right]$ and $\mathrm{E}[Y_0 \mid X, Z = z^{*}, D = 0]$ in the right-hand side of equation (5) are identified from the corresponding observed outcomes.

This implies that the mean counterfactual outcome $\mathrm{E}[Y_0 \mid X, Z = z^{**}, D = 1]$ at given X and $Z = z^{**}$ is identified from equation (5). In turn, the mean counterfactual outcome $\mathrm{E}_{X,Z=z^{**}|D=1}\mathrm{E}[Y_0 \mid X, Z = z^{**}, D = 1]$ unconditional on X and Z is identified by averaging over the observable distribution of X, Z given $D = 1$. Again, the ATT follows. Notice that if there is no selection on unobservables then the only controls used to estimate the ATT are the non-treated eligibles. In this sense, the alternative approach subsumes the instrument Z in the set of conditioning variables X.[7] With selection on unobservables, of course, the ineligible controls with $Z = z^*$ are also used in the alternative approach.

2.2 Inference

Our estimation method for the ATT closely follows the above identification proofs. For the sake of brevity we focus on the method for the main approach where we use equation (4) to obtain the mean counterfactual outcome among the treated $\mathrm{E}(Y_0 \mid D = 1)$. Equation (4) is conditional on X and $D = 0$, but we need to average it over the observable distribution of X given $D = 1$ to obtain $\mathrm{E}(Y_0 \mid D = 1)$. For this purpose we estimate a propensity score for $\mathrm{P}(D = 1 \mid X)$, using the full sample. Next, we match each treated individual to non-treated individuals, using propensity-score kernel-matching. However, contrary to the standard matching approach to treatment evaluation, we do not take the difference of the outcome of the treated and the matched (weighted mean) outcome of the controls, but we take the difference of the outcome of the treated and the matched (weighted mean) value of the right-hand side of equation (4). In the right-hand side, the separate terms are kernel-smoothed for this purpose, using propensity scores as well. The standard errors are estimated with bootstrapping. Notice that in the alternative approach based on equation (5), the

[7]See Heckman and Lozano (2004) for a discussion of the selection of covariates in matching.

main propensity score does not only depend on X but also on Z for values $Z = z^{**}$.[8] In Subsection 3.3 we discuss practical implementation issues for our estimator, in the context of the evaluation of the Swedish Youth Practice (YP) program.

We may also use the results of the previous subsection to design tests of the usual CIA assumption that $Y_0 \perp D \mid X$, if Assumptions 1 and 2 apply. The standard matching method assumes CIA, and then the first term of the right-hand side of (4) captures $\mathrm{E}(Y_0|X, D = 1)$. As already pointed out, the second term can be labelled a correction term due to selection on unobservables. Thus, the usual CIA assumption holds iff the correction term is zero for any possible X, so iff

$$\mathrm{E}\left[Y_0 \mid X, Z = z^*, D = 0\right] = \mathrm{E}\left[Y_0 \mid X, D = 0\right] \tag{6}$$

for any X. In the alternative approach (see (5)), this is replaced by

$$\mathrm{E}\left[Y_0 \mid X, Z = z^*, D = 0\right] = \mathrm{E}\left[Y_0 \mid X, Z = z^{**}, D = 0\right]. \tag{7}$$

for any possible X and z^{**}. This can again be aggregated over X and z^{**}. These equalities can be used to test the usual CIA assumption in standard matching estimation. Alternatively, we may test directly whether the correction terms are zero, because these are a by-product of the ATT estimation.

Battistin and Rettore (2008) propose a selection test based on the bias term

$$\mathrm{E}\left[Y_0 \mid X, Z, D = 0\right] - \mathrm{E}\left[Y_0 \mid X, Z, D = 1\right] \tag{8}$$

defined in regions of Z where participation is not deterministic. Under their RD design with one-sided imperfect compliance, the bias term in (8) can be computed at the

[8]The ATT estimates suggested by the two alternative approaches are not necessarily identical. This provides scope for the construction of a general specification test. However, it remains to be seen whether such a test has satisfactory power, as the underlying estimates are driven by outcomes from overlapping subsamples.

eligibility cutoff point. The statistical significance of this term at that specific point provides some information of what may happen elsewhere. In contrast, the matching setup that we explore, with arguably stronger conditions, allows to directly test the CIA on a larger part of the domain of Z, and therefore a larger sample. This is not empirically irrelevant as sample sizes often preclude meaningful analysis in local discontinuity estimation.

3 Empirical Application: Youth Practice

We study the impact of a Swedish youth employment program, the Youth Practice (YP). The aim of this program is to ease the flow of young unemployed into work by providing work experience. The main focus of our evaluation is its impact on transitions into employment. In what follows we describe the program in more detail, the data, the estimation procedure, and the results.

3.1 The program

YP is a Swedish large-scale subsidized-work program targeted at the 18-24 years old unemployed. This program was launched in July 1992. In October 1995 it was subsumed into an extended policy program for youth unemployment.

The YP program was primarily intended for unemployed individuals with a high school diploma. Participants were placed in a job in the private or public sector for 6 months with a possible extension to 12 months. In fact, eligible individuals were encouraged to find such a subsidized job themselves. While at work, YP participants received an allowance below the current wage rate. The employer paid at most a small fraction of the allowance. The job was supposed to be supplementary in the sense that it should not displace regular employment. In addition to work, participants were also expected to spend at least four hours per week at the employment office to search for more regular employment. However, the no-displacement and the job-search requirements seem to have been violated regularly (see references in Section 1).

Officially, eligibility required individuals to be younger than 25 years of age at the moment of enrolment into YP as well as to be registered with the employment office for a minimum duration of 4 months for the 20-24 years old and 8 weeks for the 18-19 years old. We restrict attention to the 20-24 years old because of a range of differences with the policy regime for those below 20 (see Forslund and Nordström Skans, 2006, and the other references in Section 1, for details on YP and youth unemployment in Sweden).

Participation was not compulsory. In fact, YP was one among several non-compulsory treatments that agents could enter. The most relevant other possible treatment is Labor Market Training, which is an expensive program that mostly consists of vocational training. But YP is by far the most common treatment among young unemployed individuals. In over 22% of the new registration spells of eligible individuals, YP is the first reported event after registration at the employment office, whereas the other possible treatments amount to only 16%, of which just over a third concerns Labor Market Training.[9]

Empirical data show that the eligibility requirement concerning the 4-month minimum registration period was not respected: almost 20 percent of participants enter YP within 1 month of registering, and over 60 percent enter before completing the first 4 months. The age eligibility rule, however, is strictly respected: participants are always below the age of 25 at the moment of enrolling into YP.

[9]We consider exits from the first unemployment period after registration with the employment office. The reported figures refer to individuals aged 20 to 24 when first registering with the employment office between July 1992 and September 1994.

3.2 Data

We use the Swedish unemployment register called *Händel*. This is an administrative dataset that comprises information from August 1991 onwards on unemployment spells, program participation and the subsequent labor market status of those who are deregistered (e.g. employment, education or inactivity). All individuals with unemployment spells since 1991 are included in the dataset and their unemployment history can be followed over time. Händel also includes demographic information on age, gender, citizenship, area of residence and education.

For the purposes of our evaluation, we use only the first registered spell starting while YP was widely available, from July 1992 to September 1994. After that, the take-up slowed substantially until YP was extinct in October 1995.[10] We use all registration spells, independently of the employment status of the new applicant.[11] We also restrict the sample to those registering as open unemployed for comparison purposes.

The Händel dataset required considerable cleaning and selection work, mainly due to the high incidence of negative and overlapping spells. The criteria applied to construct the final dataset are described in the appendix.

We take age as our variable Z, and we compare both narrow and wider age groups. The analysis is restricted to men. Table 1 reports sample sizes by eligibility and treatment status for different age groups. Each individual is represented only once in the sample as we only consider the first observed registered spell within the July 1992 to September 1994 time frame.

Column 3 in Table 1 shows that the number of program participants is small if one

[10] Among eligibles, YP occurred in only 3% of registration spells starting after September 1994 and under 1% of registration spells starting after January 1995.

[11] Employed individuals looking for a new job may register with the employment office; they account for less than 4% of all new spells for the population we are considering.

explores local variation in age to identify the impact of treatment on individuals at the eligibility threshold (row 1, column 3). This happens despite the whole population of treated being used and despite the comparatively high take-up rate among eligibles. The explanation may be a mechanical assignment issue. Although YP is the most popular treatment among young individuals in the registrar, eligibles at the verge of completing 25 years of age at inflow have a short time to enrol into the program. On the contrary, younger agents have comparatively more time, and therefore better chances, to be allocated a place. This variation in participation rates by age is shown in figure 1. It displays the rate of transition into YP by time since registration among individuals aged 24 at the moment they register and depending on whether they are at more (red curve) or less (green curve) than 4 months from completing 25 years of age. The figure shows that participation rates for the youngest cohort is steadily above zero straight from inflow, peaks at 4 months and starts declining after that. It also provides further detail to the pattern described in table 1, showing that the older cohort participation is concentrated over the first months in unemployment and is never as high as for younger cohorts. As a result, the overall hazard rates are much lower for the whole population of 24 years old at inflow (blue curve) than among those younger than 4 months from completing 25 years of age.

This pattern of participation by age creates a gradual decline in participation rates with age at inflow. Figure 2 depicts it. There is no visible discontinuity to be explored. This is not an ideal empirical setting for an application of regression discontinuity.

3.3 Estimation procedure

In this application we aim to measure the impact of treatment on the odds of leaving unemployment or finding a job some time after first registering with the employment office. The estimation uses the population of males aged close to their 25th birthday when first registering with the employment office between July 1992 and September 1994. Eligibility is based on age at inflow, where those aged 24 and below are eligible to participate in the YP and those aged 25 and above are not. Thus, age is the instrument and the 25th birthday is the cutoff point.

The *treated group* is composed of eligibles who select into YP as their first activity after registering. We consider alternative treatment groups depending on two dimensions:

1. duration of registration spell prior to enrollment into the YP: up to 3 and 6 months;

2. and distance in days to 25th birthday at registration - up to 6 months, 1 year and 2 years. [12]

Estimation of the counterfactual of interest as described in equation (4) requires two control groups. The first is the standard matching control group, drawn from the population of non-participants ($D = 0$) and reproducing the distribution of the matching variables X among the treated. Since age (Z) is not in X, non-participants comprise both non-eligible individuals and eligible individuals that opted out of YP as their first activity after registration within the considered unemployment duration. The second control group is required to compute the correction term and draws exclusively from the population of ineligibles ($D = 0, \text{age} > 24$), again reproducing the distribution of the matching variables X among the treated. In both cases, the age criterion defining alternative treatment groups, depending on distance to 25th

[12] We decided not to tighten this requirement given the small number of treated observations close to the age cutoff point (see Table 1). We also estimated the impact of treatment on the sample of individuals as far as 5 years away from their 25th birthday but the ensuing increase in the sample size causes the procedure to become forbiddingly slow when it comes to estimate the precision of the effect. It is also conceivable that our exclusion restriction does not hold for very wide age groups.

birthday at registration, is similarly applied to the construction of comparable control groups.

The alternative estimator described in equation (5) includes the instrument in the set of matching variables. In this case, the first control group, that of non-participants ($D = 0$), will be drawn exclusively from the eligibles (or those 24 or younger at inflow) who did not take up treatment as their first activity within the period of unemployment being considered.

Our matching procedure is *unconditional* on time to treatment other than through the time window requirement described above. Controlling for the duration of unemployment prior to enrolment into treatment would call for a dynamic framework which has problems of its own, in particular since we do not rule out that there is selection on unobservables; see Abbring and Van den Berg (2003, 2005). This is beyond the scope of our study. See Fredriksson and Johansson (2008) and Crépon et al. (2009), for the use of matching methods that deal explicitly with dynamic enrollment.[13]

The standard matching estimates are produced using propensity score matching with Epanechnikov kernel weights. If the instrument age is excluded from the standard matching estimates as in equation (4), the propensity score is estimated on all other observable characteristics, namely citizenship, education, region of residence, quarter of entry and labor market history during the year preceding the start of the unemployment spell. If eligibility is controlled for in the standard matching procedure as in equation (5), age in years is added to the set of covariates.

Figure 3 plots the distribution of the predicted propensity scores by treatment and eligibility status when age is excluded from the covariates set. The population being depicted is that of 24 and 25 years old at registration with the employment

[13]The severity of any resulting bias from disregarding time to treatment will depend on the time window allowed before enrolment. We therefore consider only short durations prior to enrolment into treatment, of up to 3 or 6 months, and compare results to assess for the importance of our choices.

office, where treated are individuals moving into YP in the first 6 months of the new registration spell.

Enrollment into treatment seems to be partly dependent on the observable characteristics but the distribution of the propensity score exhibits very little dependence on the eligibility status. In fact, the covariates are relatively balanced between the treated and alternative non-treated groups, even before matching, with a maximum bias of 22%. Matching on the propensity score succeeds in improving balancing for all observables, reducing the bias very substantially in most cases and to a maximum of below 4%. Results for alternative groups defined by age and/or duration of unemployment prior to treatment are very similar to these.[14]

Estimation excluded observations lying below the highest 5th percentile and above the lowest 99th percentile of the distributions of the propensity scores among treated and comparison groups. This selection procedure restricts attention to the overlapping support while moving away from the lower part of the distribution of the propensity score. Equation (4) justifies this asymmetric trimming of the distribution as the estimates of the correction term can be very imprecise for very low values of the propensity score.

3.4 Results

Our preferred sample comprises males registering with the employment office within 1 year of their 25th birthday while YP is operating in full (between July 1992 and September 1994). The sample contains 43,407 observations almost evenly split between eligibles and ineligibles (formed of individuals younger and older than 25 at entrance, respectively). Among the eligibles at inflow, just over 2% (511 observations) flow into YP within 1 month, 5.5% (1,182 observations) within 3 months and

[14]Results available from the authors under request.

almost 9% (1,887 observations) within 6 months of registering with the employment office. Our main estimates use the latter group of participants for the sake of sample size, but we will also present alternative comparisons using different age groups and unemployment durations before enrolling into YP.

Table 2 displays the estimates of the ATT on the probability of finding a regular job within 12 and 24 months of registering with the employment office.

Row 1 in the table displays the main set of estimates, based on individuals aged 24 and 25 at registration and defining treatment as flowing into YP during the first 6 months as the first destination after registration. Standard matching estimates suggest YP has a null effect on the probability of moving into employment within 12 (column 1) and 24 (column 4) months of registration. The corrected matching estimates corroborate this result when applied to the 12 months' outcome. However, the figure regarding outflows within 24 months of registration is significantly different. The correction term suggests that treated are not randomly selected once observables have been controlled for. Instead, the treated seem to be comparatively better positioned to find a job in the absence of treatment than similar non-treated. The consequence is the large and significant negative effect of treatment on outflows to employment identified by the corrected matching estimator.

To assess the robustness of this result, we tried several alternative comparisons. Some of the results are displayed in the other rows of Table 2. We restrict the sample to those registering as open unemployed in row 2. We exclude eligibles at less than 4 months of completing 25 years of age in row 3. And we restrict the control group in standard matching to be composed only of non-treated eligibles in row 4. All results are consistent with those shown in row 1. Only in row 4 are the corrected estimates after 24 months of registering not statistically significant at 5% significance level, but the exhibited pattern is similar to all other cases.

We also considered using other groups: restricting the sample to Swedish citizens;

focusing on individuals with vocational training only (the largest educational group with registration spells); and including exits to registered employment as a positive outcome. All results are consistent with the ones displayed in Table 2.[15]

The last row of Table 2 considers treatment to be 'starting YP within first 3 months after registering'. If dynamic selection issues were important at these relatively short durations, we would expect the results to show some response to such change in the definition of treatment. However, estimates in row 5 are very similar in size and pattern to those displayed in the other rows of the same table. The robustness of these results suggests our preferred time window is sufficiently narrow to keep time of treatment exogenous in this analysis.

We investigate the sensitivity of these results to age in Table 3 by varying the width of the age interval around the 25th birthday at registration.

Columns 1 to 3 display estimates of the effect of YP on the odds of finding a job within 24 months of registration. For comparison purposes, the first row repeats the last three columns in the first row of Table 2. The following two rows display results for the population of men up to 2 years (row 2) and half year (row 3) away from their 25th birthday at registration. Neither widening or narrowing the age interval changes the pattern of the results. However, results in row 3 are substantially larger but very imprecisely estimated given the small sample size.

Columns 4 to 6 display results on an alternative outcome, deregistration within 24 months of first registering. The classical matching estimate for 24-25 years old (row 1, column 4) suggests a negative overall impact of the program, maybe due to an extended lock-in effect or to the extension of eligibility to benefits as a consequence of treatment take-up.[16] A similar result holds for 23-26 (row 2). In both cases, however,

[15]Results available from the authors under request.

[16]The Swedish welfare system provides unemployment insurance for a limited amount of time after a transition from employment into unemployment. However, this period can be extended by

the correction points to the opposite direction and the resulting effect is found to be positive and statistically significant when the larger age group is used (row 2). Again here, sample size precludes a clear pattern to emerge from the analysis of the narrower age group (row 3).[17]

Results for both outcomes are considerably stable across age groups. Such lack of variation is consistent with an homogeneous effect of treatment by age for the interval being considered.

Overall, both tables suggest that standard matching may not be identifying the correct causal effect of interest (i.e., the ATT). Standard-matching results suggest that the program has no effect on the probability of finding a job and a small negative effect on the overall odds of leaving unemployment. Correcting for the potential selection bias in matching changes the picture quite substantially. The program seems to strongly reduce employment take-up in the medium run, after 24 months of registration. With regard to the overall impact on the odds of leaving unemployment, our estimation strategy suggests YP has either a zero or a small positive effect, within the same time frame. Analysis of other outcomes suggest that the possible positive effect of YP on deregistration is driven by exits into formal education (these estimates are available under request from the authors).[18]

---

participation in the programs made available by the employement offices, of which YP is one example. Repeated participation would, in principle, allow the unemployed to remain out of work and on benefits indefinitely.

[17]We have estimated effects on all outflows on other samples as in table and applying the alternative estimator as defined in equation (5). Results are similar to those discussed here.

[18]Swedish subsidized work programs have been the focus of other studies. In particular, Sianesi (2004) analyzes the overall impact of the Swedish ALMP system and the differential impact of each of the numerous available treatments for adults (so this excludes YP). She finds that subsidized employment is the best performer in terms of moving unemployed back into work, and that the positive effect of subsidized employment seems to last. All other programs have either a zero or a negative impact, possibly arising through the renewed eligibility to benefits as a consequence of

4 Conclusion

We have developed and applied an evaluation method for the effects of program participation (or policy exposure) on individual outcomes, if participation is selective but individuals are ineligible in case of a certain value of some observed instrumental variable. From a practical point of view this is a common setting, in particular for active labor market policies for young individuals. In those cases, participation may be selective because individuals can choose between different programs and/or because the duration until enrollment is not deterministically set. Program participation is only possible if the individual is aged below a certain age. With selective participation, if the CIA is violated, matching cannot be used. For the same reason, one cannot simply compare those below the threshold who are treated to those above the threshold (who are all non-treated). However, our novel method, which exploits the eligibility boundary restriction within the matching framework, provides consistent estimates of the average treatment effect on those who are treated.

Our approach relies critically on the availability of an instrument satisfying Assumptions 1 and 2 in Section 2. Assumption 2 is automatically satisfied in our preferred practical application of a policy that allows for selective participation only

program participation. Larsson (2003) studies specifically the effects of YP on exits to employment and finds negative effects 12 months after treatment using standard matching techniques. More generally, youth programs have often shown disappointing results. Heckman, LaLonde and Smith (1999) survey a large number of evaluation studies on US and European programs with more negative results in the US than in Europe. More recently, the survey by Bergemann and Van den Berg (2008) of evaluation results in Europe by gender finds that young men do not seem to benefit from these interventions in terms of labor market outcomes, while young women are found to benefit more frequently. A noticeable exception are programs that mix improved job-search assistance and tougher job-search monitoring such as the British New Deal for Young People. This type of programs has shown more consistent positive effects (e.g. Blundell et al., 2004, De Giorgi, 2005, Anderson, 2000, Van den Berg et al., 2004, and Van den Berg and Van der Klaauw, 2006).

on certain values of an observed variable. To obtain precise estimates of our correction term, however, we also require the program to generate a reasonable number of participants to avoid dividing by a number close to zero.

The application to the Swedish Youth Practice program shows that our method can deliver evaluation results that differ from those based on standard matching methods. The standard matching estimates for the effect on re-employment are always zero, whereas the estimates based on our method can be strongly negative. The difference between the estimates is systematically significant when the outcome of interest is "finding a job within 24 months of becoming unemployed". The effects on the overall exit probabilities out of unemployment are invariably estimated to be smaller than those based on matching, although the differences here are not significant. As a result, we are more pessimistic about the effect of subsidized work on the rate of finding work than if we had incorrectly based ourselves on the matching estimates, while overall exit rates from the registrar are less negative (even possibly positive) than if we had relied on standard matching alone. The latter are driven by outflows into formal education. From a policy point of view, our results suggest that perhaps the optimism about the use of subsidized work programs to bring unemployed youth back to work should be tempered.

References

Abbring, J.H. and G.J. van den Berg (2003), "The non-parametric identification of treatment effects in duration models", *Econometrica*, 71, 1491–1517.

Abbring, J.H. and G.J. van den Berg (2005), "Social experiments and instrumental variables with duration outcomes", Working paper, IZA, Bonn.

Anderson, P. (2000), "Monitoring and Assisting Active Job Search", OECD Proceedings, *Labour Market Policies and the Public Employment Service*

Andrews, D. and M. Schafgans (1998), “Semiparametric Estimation of the Intercept of a Sample Selection Model,” *Review of Economic Studies*, 65, 497–517

Battistin, E. and E. Rettore (2008), “Ineligibles and Eligible Non-Participants as a Double Comparison Group in Regression Discontinuity Designs”, *Journal of Econometrics*, 142, 715–730

Bergemann, A.H. and G.J. van den Berg (2008), “Active labor market policy effects for women in Europe – a survey”, *Annales d’Économie et de Statistique*, 91/92, 385–408

Bijwaard, G. and G. Ridder (2005), “Correcting for selective compliance in a re-employment bonus experiment”, *Journal of Econometrics*, 125, 77–111.

Blundell, R., M. Costa Dias, C. Meghir and J. Van Reenen (2004), “Evaluating the Employment Impact of a Mandatory Job Search Program”, *Journal of the European Economic Association*, 2(4), 569–606

Bonnal, L., D. Fougère, and A. Sérandon (1997), “Evaluating the Impact of French Employment Policies on Individual Labour Market Histories,” *Review of Economic Studies*, 64, 683–713

Burgess, S., C. Propper, H. Rees and A. Shearer (2003), “The Class of 1981: The Effects of Early Career Unemployment on Subsequent Unemployment Experiences”, *Labour Economics*, 10, 291–309

Card, D. and D. Sullivan (1988), “Measuring the Effect of Subsidized Training Programs on Movements in and out of Employment,” *Econometrica*, 56, 497–530

Cochrane, W. and D. Rubin (1973), “Controlling Bias in Observational Studies,” *Sankyha*, 35, 417–446

Crépon, B., M. Ferracci, G. Jolivet and G.J. van den Berg (2009), "Active labor market policy effects in a dynamic setting", *Journal of the European Economic Association*, 7, 595–605

De Giorgi, G. (2005), "Long-term effects of a mandatory multistage program: the New Deal for Young People in the UK", Working paper, IFS, London

Forslund, A. and O. Nordström Skans (2006), "Swedish youth labour market policies revisited", Working paper, IFAU, Uppsala

Fredriksson P. and P. Johansson (2008), "Dynamic Treatment Assignment - The Consequences for Evaluations using Observational Data", *Journal of Business and Economics Statistics*, 26, 435–455.

Gritz, R. (1993), "The Impact of Training on the Frequency and Duration of Employment," *Journal of Econometrics*, 57, 21–51

Heckman, J., H. Ichimura, and P. Todd, (1998) "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, 65, 261–294

Heckman, J., H. Ichimura, J. Smith, and P. Todd (1998) "Characterization of Selection Bias Using Experimental Data," *Econometrica*, 66, 1017–1098

Heckman, J., R. LaLonde and J. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs" in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Volume 3, North-Holland, Amsterdam

Heckman, J., and S. Lozano (2004), "Using matching, instrumental variables and control functions to estimate economic choice models", *Review of Economics and Statistics*, 86, 30–57

Heckman, J. and R. Robb (1985) "Alternative Methods for Evaluating the Impact of Interventions," in J. Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data*, Cambridge University Press, New York

Kluve, J. (2006), "The Effectiveness of European Active Labor Market Policy", Working paper, RIW, Essen

Larsson, L. (2003), "Evaluation of Swedish Youth Labor Market Programs", *Journal of Human Resources*, 38, 891–927

Richardson, K. and G.J. van den Berg (2001), "The Effect of Vocational Employment Training on the Individual Transition Rate from Unemployment to Work," *Swedish Economic Policy Review*, 8, 175–213

Rosenbaum, P. and D. Rubin (1983) "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41–55

Sianesi, B. (2004), "An Evaluation of the Swedish System of Active Labour Market Programmes in the 1990s", *Review of Economics and Statistics*, 86, 1, 133–155

Van den Berg, G.J., B. van der Klaauw, and J.C. van Ours (2004), "Punitive Sanctions and the Transition Rate from Welfare to Work," *Journal of Labor Economics*, 22, 211–241

Van den Berg, G.J. and B. van der Klaauw (2006), "Counselling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment, *International Economic Review*, 47, 895–936

White, M. and G. Knight (2002), "Benchmarking the effectiveness of NDYP: A review of European and US literature on the microeconomic effects of labour market programmes for young people", Working paper, PSI, London

Appendix: Data cleaning and selection

Händel is an administrative dataset comprising information on all registered unemployment spells from August 1991 onwards. It details longitudinal information on the whole population of registered spells, including any undertaken treatments, the history of earned subsidies, destination on leaving the registrar and demographics such as age, citizenship, education and usual occupation.

The main obstacle in using Händel is the frequency of negative and overlapping spells. We have dealt with these occurrences in a conservative way to minimize any resulting bias introduced by data handling.

To start with, we created a condensed variable describing labor market status while in registrar. The four broad categories considered are: unemployment, registered employment, YP, all other possible treatments. Using these, we collapsed all overlapping spells in the same broad category. Spells in different broad categories overlapping by no more than 2 weeks were corrected by setting the exit date of the earliest equal to the entry date of the latest as exit dates are generally more imprecise. Zero duration spells were discarded. At last, individual histories with a remaining error were censored from the time of the error onwards and a censoring indicator was created to correct estimates for the possibility of censored histories.

Data selection followed a number of criteria. First, we used only males. Then we selected individuals starting a new registered *unemployment* spell during the period YP was more popular, between July 1992 and September 1994. Of all the selected spells, we kept only the first one and followed the corresponding individuals over time to find out about treatment take up and labour market outcomes. We considered individuals aged between 20 and 29 at the time of registration and classified as eligibles those aged 24 or younger.

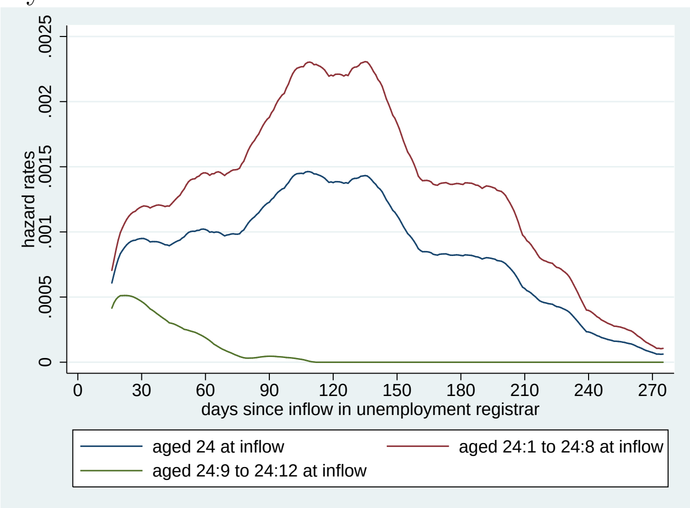
In running the estimation procedure, we also focused on more narrowly defined groups as defined by the following variables: (*i*) Distance to 25th birthday at inflow; (*ii*) Nationality; (*iii*) Employment status at registration; and (*iv*) Education attainment. These alternative comparisons are specified in the main text.

Table 1: Number of observations by age group and eligibility/treatment status; age groups defined by distance to 25th birthday when first registering with the employment office between July 1992 and September 1994; men only.

Distance to 25th birthday at inflow		ineligibles (over 25)	eligibles (under 25)		Total
			non-participants	participants	
		(1)	(2)	(3)	(4)
(1)	up to 3 months	5,444	5,240	81	10,765
(2)	up to 1 year	21,950	19,428	2,029	43,407
(3)	up to 2 years	43,683	37,118	6,064	86,865
(4)	up to 5 years	102,450	112,501	32,528	247,479

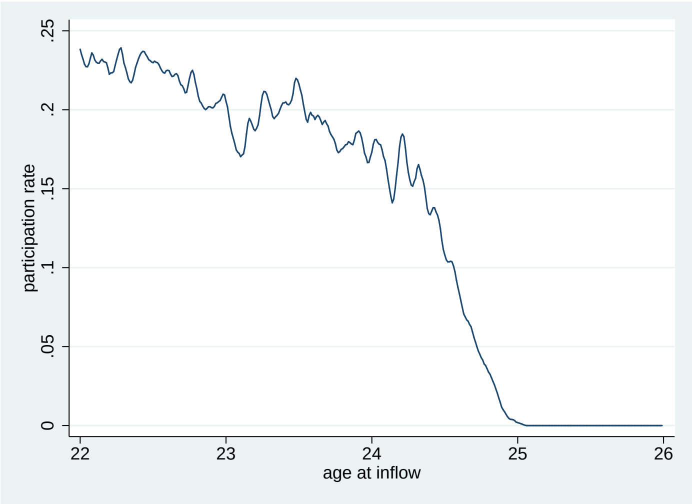
Notes: Population of males close to the 25th birthday when registering with the employment office between July 1992 and September 1994. Eligibles (ineligibles) are those aged 24 and below (25 and above) at registration. Participants are those taking YP as the first event after registration.

Figure 1: Hazard rates into YP by duration of unemployment spell and age at registration with employment office



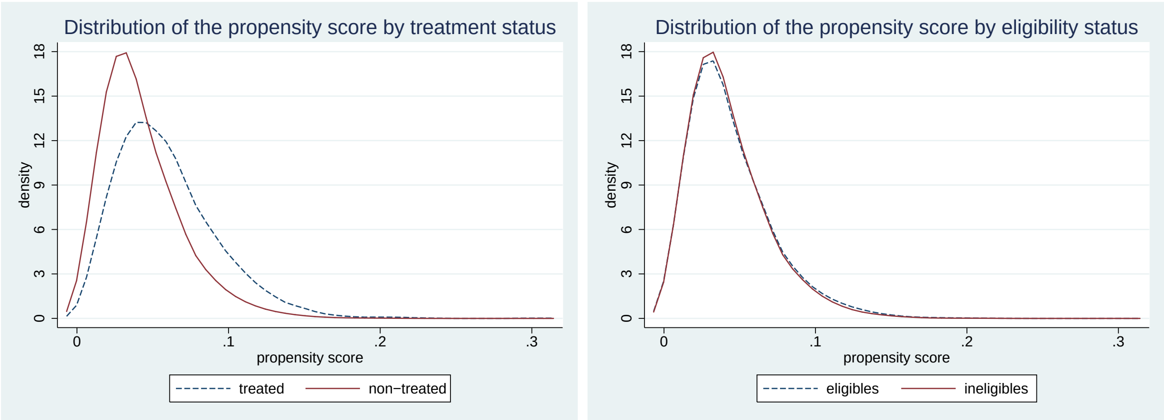
Notes: Plotted curves are smoothed Kaplan Meyer hazard rates using Local Linear Regression with a bandwidth of 15 days. Population of males aged 24 when registering with employment office between July 1992 and September 1994.

Figure 2: Probability of participation by age at inflow into new registered unemployment spell; men only



Notes: Population of males aged 22 to 26 when registering with the employment office between July 1992 and September 1994. "Participation" means flowing into YP as first event after registration.

Figure 3: Probability density functions: propensity score



Notes: Plotted curves are probability density functions of propensity scores estimated on the population of men aged 24 and 25 when registering with the employment office between July 1992 and September 1994. Treated are 24 years old moving into YP as first destination within 6 months of inflow. Non-treated are 24 and 25 years old not participating in YP as first event within 6 months of inflow. Eligibles (ineligibles) are those aged 24 (25) at registration.

Table 2: ATT on the outflows to regular employment

	Outcome: finding a job						
	within 12 months of registering			within 24 months of registering			
	classical matching	correction term	adjusted matching	classical matching	correction term	adjusted matching	nr of observ.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Treated: 24 years old moving into YP within 6 months of registration *Controls:* ineligibles (25 years old) and eligible (24 years old) non-treated						
(1)	-0.009	-0.015	0.006	-0.008	0.096*	-0.104*	1,699
	(0.011)	(0.045)	(0.047)	(0.012)	(0.048)	(0.050)	
	Treated: 24 years old moving into YP within 6 months of registration - open unemployed *Controls:* ineligibles (25 years old) and eligible (24 years old) non-treated - open unemployed						
(2)	-0.012	-0.022	0.010	-0.010	0.100*	-0.109*	1,606
	(0.011)	(0.045)	(0.048)	(0.013)	(0.050)	(0.051)	
	Treated: 24:1 to 24:8 years old moving into YP within 6 months of registration *Controls:* ineligibles (25 years old) and eligible (24:1 to 24:8 years old) non-treated						
(3)	-0.011	-0.002	-0.008	-0.004	0.070*	-0.073*	1,579
	(0.012)	(0.036)	(0.038)	(0.014)	(0.035)	(0.036)	
	Treated: 24 years old moving into YP within 6 months of registration *Controls:* eligible (24 years old) non-treated						
(4)	-0.015	-0.041	0.026	-0.007	0.042	-0.049	1,563
	(0.013)	(0.048)	(0.047)	(0.014)	(0.052)	(0.051)	
	Treated: 24 years old moving into YP within 3 months of registration *Controls:* ineligibles (25 years old) and eligible (24 years old) non-treated						
(5)	0.006	0.005	0.001	0.004	0.154*	-0.150*	1,049
	(0.014)	(0.070)	(0.071)	(0.017)	(0.075)	(0.075)	

Notes: Estimates for males only. Sample selection criteria varies by row as detailed in row titles. All estimates based on sample of new registrations with the employment office. "Treatment" in rows 1 to 5 stands for flowing into YP within 6 months of registering with employment office as first destination after registration. Row 1 compares treated aged 24 at registration with non-treated aged 24 or 25 at registration. Row 2 restricts the sample to those registering as open unemployed. Row 3 restricts the sample of eligibles to 24 years old at more than 4 months from their 25th birthday at registration. Row 4 restricts the control group in standard matching to the eligibles (aged 24 at registration). Finally, row 5 redefines "treatment" as flowing into YP as first destination within 3 months of registration and compares treated aged 24 at inflow with non-treated aged 24 or 25. The impact of treatment is estimated on the probability of moving into employment within 12 months (columns 1 to 3) and 24 months (columns 4 to 6) of registration. Columns 1 and 4 display standard matching estimates. Columns 2 and 5 display the correction term as specified in the right-hand side of equation (4) or, for row 4, of equation (5). Columns 3 and 6 display the corrected matching estimates using the counterfactuals as specified in equation (4) or, for row 4, in equation (5). Matching on the propensity score using kernel Epanechnikov weights with a bandwidth of 0.02 for a probability ranging in the unit interval. Bootstrapped standard errors based on 200 replications in brackets below the estimate.

* Statistically different from zero at 5% significance level.

Table 3: ATT on outflows to employment and deregistration within 24 months of registration

	Outcome: employment			Outcome: deregistration			
	classical matching	correction term	adjusted matching	classical matching	correction term	adjusted matching	nr of observ.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated: 24 years old moving into YP within 6 months of registration *Controls:* ineligibles (25 years old) and eligible (24 years old) non-treated							
(1)	-0.008	0.096*	-0.104*	-0.032*	-0.075	0.043	1,699
	(0.012)	(0.048)	(0.050)	(0.012)	(0.049)	(0.051)	
Treated: 23-24 years old moving into YP within 6 months of registration *Controls:* ineligibles (25-26 years old) and eligible (23-24 years old) non-treated							
(2)	0.006	0.082*	-0.076*	-0.031*	-0.151*	0.120*	4,468
	(0.008)	(0.028)	(0.028)	(0.008)	(0.030)	(0.030)	
Treated: 24:7 to 24:12 years old moving into YP within 6 months of registration *Controls:* ineligibles (25:1 to 25:6 years old) and eligible (24:7 to 24:12 years old) non-treated							
(3)	-0.001	0.187	-0.188	-0.011	0.098	-0.109	401
	(0.026)	(0.152)	(0.154)	(0.026)	(0.147)	(0.150)	

Notes: Estimates for males only. Sample selection criteria varies by row as detailed in row titles. All estimates based on sample of new registrations with the employment office. “Treatment” stands for flowing into YP within 6 months of registering with employment office as first destination after registration. Row 1 compares treated aged 24 at registration with non-treated aged 24 or 25 at registration. Row 2 uses the sample of individuals at less than 2 years from 25th birthday on registration. Row 3 uses only individuals at less than 6 months from their 25th birthday at registration. The impact of treatment is estimated on the probability of moving into employment within 24 months (columns 1 to 3) and deregistering within 24 months (columns 4 to 6) of inflow. Columns 1 and 4 display standard matching estimates. Columns 2 and 5 display the correction term as specified in the right-hand side of equation (4). Columns 3 and 6 display the corrected matching estimates using the counterfactuals as specified in equation (4). Matching on the propensity score using kernel Epanechnikov weights with a bandwidth of 0.02 for a probability ranging in the unit interval. Bootstrapped standard errors based on 200 replications in brackets below the estimate.

* Statistically different from zero at 5% significance level.