n° 2008-25

Active Labor Market Policy Effects in a Dynamic Setting

B. CRÉPON[1], M. FERRACCI[2], G. JOLIVET[3], G. van den BERG[4]

[1] CREST-INSEE, CEPR and IZA.
[2] University of Marne-la-Vallée and CREST-INSEE.
[3] University of Bristol.
[4] VU University d'Amsterdam, IFAU-Uppsala, IZA, Netspar and CEPR.

Active Labor Market Policy Effects in a Dynamic Setting

**Bruno CRÉPON, Marc FERRACCI, Grégory JOLIVET,
Gerard J. van den BERG**

September 2008

Abstract

This paper develops and implements a method to identify and estimate treatment effects in a dynamic setting where treatments may occur at any point in time. By combining the standard matching approach to the timing-of-events approach, it demonstrates that the effect of the treatment on the treated at a given date can be identified although non-treated may be treated later in time. The approach builds on a "no anticipation" assumption and the assumption of conditional independence between the duration before treatment and the duration before exit. To illustrate the approach, the paper studies the effect of training for unemployed workers in France, using a rich register data set. Training has little impact on unemployment duration. The contamination of the standard matching estimator due to later entries into treatment is large if the treatment probability is high.

Résumé

Cet article développe et applique une méthode permettant d'identifier et d'estimer les effets d'un traitement dans un cadre dynamique, où le traitement peut intervenir à différents points du temps. En reliant la méthode du "matching" à celle du "timing-of-events", nous montrons que les effets du traitement sur les individus traités à une date donnée peuvent être identifiés, même si certains individus du groupe de contrôle sont traités ultérieurement. La méthode repose sur l'hypothèse de non anticipation du traitement, et sur l'hypothèse d'indépendance conditionnelle entre la durée avant traitement et la durée de chômage. Pour illustrer cette approche, nous étudions les effets de la formation des chômeurs en France, à l'aide d'une riche base de données administrative. La formation a peu d'impact sur la durée de chômage. Il apparaît que les entrées ultérieures en traitement produisent une contamination de l'approche par appariement standard d'autant plus importante que la probabilité de traitement est élevée.

1   Introduction

In terms of their expenses, active labor market policy (ALMP) programs are the most important policy measures aimed at increasing the outflow from unemployment to work. These programs include job search assistance, training, subsidized work, and their combinations (see e.g. Carcillo and Grubb, 2006, for a recent overview). A typical feature of ALMP is that program participation is not instantaneous upon inflow into unemployment. Instead, individuals are observed to enter ALMP programs at any possible elapsed unemployment duration even though participation is not prohibited by entitlement restrictions. This reflects the assignment process. Case workers are reluctant to assign workers too early, because many of them re-enter employment relatively fast anyway. The starting date of a training program may depend on whether a sufficient number of potential trainees is available. Likewise, this date may be delayed by quantity constraints. The availability of subsidized work depends on the inherent randomness in the moment at which vacancies are created. More in general, the effort levels of the unemployed worker and his case worker may display random fluctuations over time.

The variation in the timing of program participation (or, shortly, the treatment date) means that those who are not treated at a given elapsed unemployment duration, say t_s, may be treated later. As has been recognized in the literature (see e.g. Fredriksson and Johansson, 2008), this has an important methodological implication. The difference between the residual unemployment durations of the treated at t_s and the non-treated at t_s partly reflects the effect of later treatments on the durations of those who are not yet treated at t_s. Application of the standard (propensity score) matching method may then lead to biased outcomes. This is unfortunate in the light of the attractive other features of matching as an evaluation method for average treatment effects. In particular, under the assumption of conditional independence (CIA) of potential outcomes and assigned treatment conditional on covariates, matching is particularly well-equipped for evaluation in the presence of effect heterogeneity. Notice that the methodological complication cannot be solved by discarding outcomes of non-treated who are observed to be treated later and retaining outcomes of non-treated who are observed not to be treated before they exit unemployment. After all, such an approach produces samples that are selective in terms of the outcomes.

Abbring and Van den Berg (2003, 2004) point out that under the conditional independence assumption (no-selectivity) and the assumption that future treatments do not affect current outcomes (no-anticipation), the natural control group for the treated at t_s is formed by the not-yet treated at t_s. Comparing the treated and not-yet treated at t_s straightforwardly identifies the average *instantaneous* effect on the transition rate from unemployment to work at t_s.

This is because any later treatments among the not-yet treated at t_s are not informative on the outcomes among the not-yet treated *at t_s*. Fredriksson and Johansson (2008) apply this idea in a discrete-time setting. More importantly, they develop a matching estimator for the average effect on non-instantaneous outcomes, notably the residual unemployment duration. For given t_s, this estimator compares outcomes for the treated and not-yet treated at t_s, where the outcomes of the latter are only used insofar as they remain not-yet treated. In general, estimation of standard errors is left open.

In this paper we clarify and advance on the literature. We develop a comprehensive matching evaluation framework in continuous time with counterfactual duration outcomes. By assuming that the dynamic assignment process is driven by a single index, it follows that the propensity score is captured by the systematic (single-index) part of the hazard rate of the duration until treatment. This can be conveniently estimated. Next, average treatment effects on the treated can be estimated with matching, given CIA and the no-anticipation assumption. The estimator is similar to the Fredriksson and Johansson (2008) estimator, but we derive and motivate it from the point of view of the well-known counterfactual duration framework of Abbring and Van den Berg (2003). Moreover, standard errors for average effect estimates based on kernel matching or inverse probability weighting are obtained by bootstrapping (see Heckman, Ichimura and Todd, 1998, and Hirano, Imbens and Ridder, 2003).

We apply our estimator to study the effect of participation in a training program on the unemployment duration distribution in France. The data are informative on individual past labor market outcomes including past ALMP participation. We estimate average effects and we analyze the contamination bias of the standard approach.

Although the paper is written as dealing with ALMP evaluation, it is clear that our methodological approach is not tied to that. We should point out that there are alternative approaches to deal with the above methodological issue. One may proceed to estimate semi-parametric models that model outcomes and treatments at various points in time and that include systematic heterogeneity into the model, to obtain estimates of the treatment effect as a function of the time elapsed since enrollment, the elapsed unemployment duration, and heterogeneity indicators (see Richardson and Van den Berg, 2008). Yet another approach is to apply sequential CIA at each point of time conditional on events that took place earlier (Lechner and Miquel, 2005). We should also note that an alternative to comparing treated and not-yet treated is to compare those who enter one type of ALMP at t_s to those who enter a different type of ALMP at t_s. If the two ALMP are mutually exclusive in the sense that an individual cannot participate sequentially in both, then such a comparison is informative on the relative merits of the two treatments (see e.g. Imbens, 2000, and Lechner, 2002).

3

2 Identification of dynamic treatment effects

2.1 The model

Consider a group of individuals who enter a given state U (say unemployment) at date 0 and let T_u be the duration spent in U. Assume that a treatment is available at a random date T_s. We are interested in the effect of receiving the treatment at date t_s. To define this effect, we consider the *potential* durations $T_u(t_s)$ spent in U when treatment occurs at date t_s.[1] In our framework T_s is a latent variable as it can be censored by T_u. To keep the analysis simple, we assume time is discrete.

To evaluate the effect of the treatment, we need a counterfactual. Ideally, we would like to define a duration associated to "not being treated" but, without further assumptions, this is a vague notion since the data typically do not contain this information. We know that individuals have not been treated up to a certain date but if they leave without treatment we do not know what would have happened later on. More formally, if an individual leaves at date t without having been treated, his duration could have been ruled by a process $T_u(t')$ where $t' > t$. To avoid this issue, we make a "no anticipation" assumption similar to the one in Abbring and van den Berg (2003) and in the evaluation studies based on the "timing of events" approach:

$$P\left(T_u(t') = t\right) = P\left(T_u(t'') = t\right), \quad \forall t < \min(t', t''). \tag{A1}$$

This assumption means that for each individual, all processes have the same distribution prior to the treatment dates. This common distribution then characterizes a new process denoted as $T_u(\infty)$:

$$P\left(T_u(\infty) = t\right) = P\left(T_u(t') = t\right), \quad \forall t < t'. \tag{1}$$

$T_u(\infty)$ is the duration of an individual who, at any date, has the same probability of leaving than those who will be treated later. We will use this duration as our counterfactual.

Now consider a function G and a treatment date t_s. We are interested in parameters of the form:

$$TT_G(t_s) = E\left[G\left(T_u(t_s)\right) - G\left(T_u(\infty)\right) \big| T_s = t_s, T_u(\infty) > T_s\right]. \tag{2}$$

We thus focus on individuals entering treatment at date t_s and who have not yet left at this date.[2]

---

[1]We assume that treatment is instantaneous and do not model the duration in treatment.

[2]Note that we can replace $T_s = t_s, T_u(\infty) > T_s$ with $T_s = t_s, T_u(t_s) > t_s$ in the conditionning thanks to assumption (A1). For the same reason, there is little interest in looking at the difference for the treated before the date of treatment as the two processes are similar.

In this paper, we will consider the following set of G functions: $G_\tau(t) = 1\{t > \tau + t_s\}$, $\tau > 0$. Then, the $TT_G(t_s)$ parameters give the difference between the survivor functions, at date $t_s + \tau$, of individuals treated at date t_s and those at risk[3] but not yet treated at date $t_s + \tau$.[4]

We also consider more aggregated parameters, defined as treatment effects on individuals entering treatment within a given period:

$$TT_G(t_{s1}, t_{s2}) = E\left[G\left(T_u(T_s)\right) - G\left(T_u(\infty)\right) \big| t_{s1} \leq T_s \leq t_{s2}, T_u(\infty) > T_s\right] \quad (3)$$

2.2 Identification

The main targets of identification are the parameters $TT_{G(t_s)}$ defined by (2). These differences involve two quantities. The first one is the average outcome "if treated for the treated":

$$E\left[G\left(T_u(t_s)\right) \big| T_s = t_s, T_u(\infty) > t_s\right] \quad (4)$$

We can identify this term using the "no anticipation" assumption (A1) to replace $T_u(\infty) > t_s$ with $T_u(t_s) > t_s$ and thus to consider individuals who have actually been treated at date t_s. Note that when durations up to exit from state U are censored, assumption (A1) is not enough. We will introduce later additional assumptions allowing to estimate our parameter even in the case of censored durations in state U.

The second term of the difference in (2) is the conditional average counterfactual:

$$E\left[G\left(T_u(\infty)\right) \big| T_s = t_s, T_u(\infty) > t_s\right]. \quad (5)$$

Its identification raises the standard issue that the counterfactual is not observed for the treated. We adapt the approach found in most of the matching literature and assume that we observe a set of individual characteristics X such that, conditional on X, assignment to treatment is independent of the counterfactual. Our assumption is different however from what is done in the usual case as the independence is assumed between the *latent* variables $T_u(\infty)$ and T_s:

$$T_u(\infty) \quad \perp \quad T_s \quad | \quad X. \quad (A2)$$

The two assumptions (A1) and (A2) allow us to identify the missing term (5). Identification is based on two important steps.

---

[3]Individuals at risk are those who have not yet left the state of interest. For example, in our application individuals at risk at date t are those who are still unemployed at t.

[4]Note that, since the TT_G parameters we consider involve survival functions, their sign leads to a different interpretation from what is usually found in the evaluation literature. Indeed, if $TT_G(t_s)$ is positive, treating individuals at date t_s decreases their probability of leaving before $t_s + \tau$. If, as will be the case in our empirical application, one is interested in the effect of training on unemployment, a positive TT_G parameter indicates that training individuals tends to lengthen their unemployment spells.

First, assumption (A2) suggests to consider individuals who are at risk and not yet treated at t_s as a potential control group:

$$E\left[G\left(T_u(\infty)\right)\big|T_s = t_s, T_u(\infty) > t_s, X\right] = E\left[G\left(T_u(\infty)\right)\big|T_s > t_s, T_u(\infty) > t_s, X\right].$$

Then, for individuals in the control group, we observe either $T_u(\infty)$ if individuals leave U before entering treatment, or $T_u(\infty)$ censored by T_s if individuals enter treatment before leaving U. The important point is that because of assumption (A2), conditionally, on X this censoring is independent from $T_u(\infty)$. It is well known that this information is enough to identify the distribution of $T_u(\infty)$ and therefore our parameter.

It may be the case that the duration in state U is not always observed. This duration can indeed be censored. For example when U is unemployment, durations are usually censored if the unemployment spells exceed the date up to which information is available. To deal with this issue, we make another conditional independence assumption. If T_C denotes the durations before censoring, we assume that:

$$T_C \quad \perp \quad (T_u(t_s), T_s) \quad | \quad X, \quad \forall t_s. \tag{A3}$$

Under assumption (A3), the distribution of $T_u(t_s)$ conditional on being treated at date t_s and X is identified even in the case of censored durations. Therefore, the two terms of the parameters $TT_G(t_s)$ of interest are both identified.

2.3 Estimation

We present a simple way to estimate our parameters of interest. We concentrate on the effect on the survival function, but every other parameter corresponding to equation (2) could be estimated. The basic idea is to match treated and non treated populations with the same value of X. In the standard matching literature, it is well known that it is possible to deal with the dimensionality of the X vector by using the propensity score property (cf. Rosenbaum and Rubin, 1983). Such a property can be extended to the case of multi-valued treatment as follows:[5] assuming that there is an index $s(X)$ such that $f_{T_s|X}(t) = f(t, s(X))$, then the CIA (A2) implies $T_u(\infty) \quad \perp \quad T_s \quad | \quad s(X)$.

Our estimation method consists of two steps. In the first step we estimate the propensity score. Note that the duration up to treatment is observed for individuals entering treatment and censored for individuals leaving unemployment before treatment. Therefore the estimation of the score just involves information about T_s.

---

[5]See Crépon and Deplatz (2001) and Hirano and Imbens (2004).

For a given treatment date t_s, the treatment group consists of individuals still unemployed and entering treatment at t_s. The potential control group consists of individuals still unemployed at this date but not yet treated.

Then, we proceed to the matching step. We match individuals on the score $s(X)$. Initial populations of treated and non treated can be split into subpopulations with close values of the score.[6] This procedure departs from the matching literature (e.g. Sianesi, 2004) and is closer to blocking methods (see Cochran, 1968, or Rosenbaum and Rubin, 1983). In each cell, the hazard function h_t of the residual duration in state U is estimated. It is simply computed as the number of individuals leaving unemployment at t (with non censored duration) divided by the number of unemployed still at risk at this date.[7] That is, individuals still unemployed at t for the control group and individuals still unemployed at t *but not yet treated* for the control group. From the hazard function we can derive the survival function $\overline{F}$ as $\overline{F}(t) = (\tau) = \prod_{t=t_s}^{t_s+\tau} (1 - h_t)$. Once the survival functions have been estimated, the difference between treated and non treated is averaged using the distribution of the score function in the treatment population.

3   An empirical illustration using active labor market policies in France.

3.1   Data and specifications

We apply our method to the evaluation of training programs for unemployed workers in France. See Crépon, Ferracci and Fougère (2007) for a description of the French unemployment insurance and training systems. We consider as treatment any first entry into any training program, and we are interested in the effect of this treatment on unemployment duration.

We use data from the Fichier National des Assedic (FNA) which is the national register of all unemployed workers in France since 1990. Each quarter, a random 2.5% sample is drawn from this register. Our data set consists of the four draws made in 2007. We observe all the unemployed spells of each individual in our sample from 1990 to March 2007. For the analysis, we consider all unemployment spells starting between 2002 and 2004. We start only in 2002 because of a major reform which took place in Autumn 2001. We do not consider spells starting in January 2005 and after because we want to limit the exogenous censoring due to the draw date. We end up with 201 277 spells, 6.4% of which have not ended in March 2007 and are thus censored.

---

[6]For example, we may consider the population defined by the percentiles of the distribution of the score in the treated population.

[7]With propensity score matching, we have to strengthen (A3) and assume that $T_C \perp T_u(t_s)|T_s, X$.

We include a rich set of covariates in the X vector to ensure that the CIA (A2) holds. These are the following: age, gender, occupation of the previous job (7 categories), region (23 regions), duration of affiliation to the unemployment insurance system, unemployment benefits, wage in the previous job and a dummy equal to one if the occupation in the job searched is the same as the one in the previous job. In addition to these controls, we use the longitudinal dimension of our data to control for individual unemployment and training histories.

The first step is the estimation of the propensity score. We consider a simple proportional hazard model and leave more elaborate estimators to future research. The duration dependence of the hazard function is chosen as a piecewise constant function. We allow for 11 cutting points regularly distributed over the interval $[0, 18$ months$]$. Heterogeneity is captured by the single-index $X\beta$. We also add an additional unobserved heterogeneity term, which is modelled as a multiplicative binary variable. The score function $\widehat{s}(X)$ is simply the product $X\widehat{\beta}$.

For the second step, we define cells in the treatment and control groups based on the 24 percentiles of the distribution of the score in the treatment group. We then proceed as explained in Subsection 2.3.

3.2   Results

We first consider basic estimators for various treatment dates. The time unit is the month. The upper panel of Figure 1 shows the treatment effect on the treated (remember that a positive effect means longer unemployment). The dashed lines delimit the confidence interval obtained by bootstrapping. We consider two treatment dates: $t_s = 3$ and $t_s = 9$.

The figures clearly show that the effect at these two treatment dates have the same pattern. There is first a locking in period that lasts around 18 months. The effect is important as the survival rate can increase by 10%. After this locking in period, the effect of training on survival is negative. However the effect is small (around 2% three years after entry) and significant only for $t_s = 3$. The overall picture is therefore close to what as already pointed out in numerous studies: training the unemployed does substantially shortens their unemployment spells.

We also estimate more aggregated parameters such as (3) for treatment dates ranging from 1 to 6 and from 7 to 12. To do this we estimate the basic parameters for each date and we average them using the distribution of the treatment date. The lower panel of Figure 1 presents results for the effect of treatment on the treated either when treatment starts within the six first months of the unemployment spell or when it starts within the next six months. Implementing these aggregated parameters does not change the overall picture about the effect of the policy. Training reduces the survival rate on unemployment only

in the long run and the effect is small. There are sizable efficiency gains to consider aggregated parameters. These gains are sufficiently large for the effect in the long run to be significant when entry into treatments occurs late in the unemployment spell.

Figure 1: Estimated treatment effect on the treated on the survival function
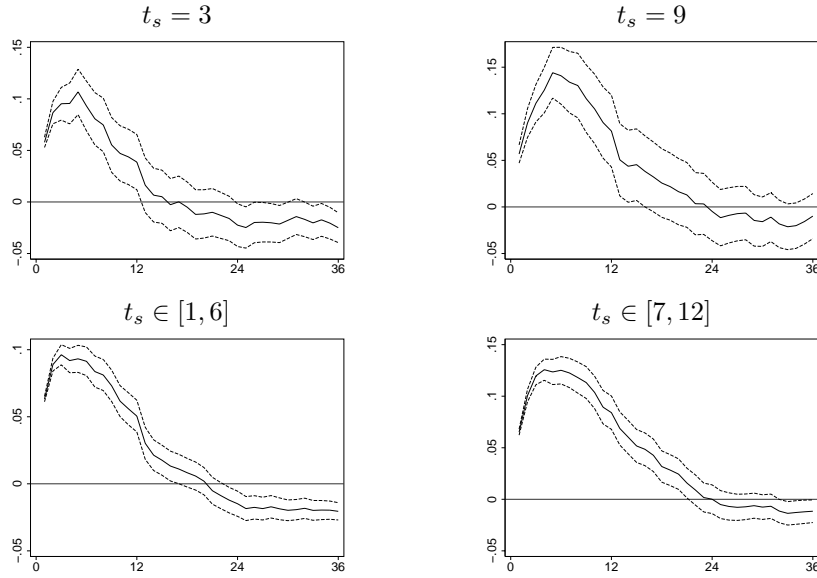
There are two main differences between our dynamic matching method and the standard matching approach. The first one is related to exits: in our setting we only count exits without censoring, while the standard matching method also includes them. More importantly, the second difference lies in the definition of the risk sets (and thus of exits). To fix ideas, let $R_\tau^S(t_s)$ be the risk set at date $\tau + t_s$ for individuals with treatment status $S = 0, 1$, for the treatment date t_s. Let $X_\tau^S(t_s)$ be the corresponding set of exits from unemployment (so that the hazard rate can be estimated as X/R) and assume exogenous censoring away for simplicity. We have:

<table>
<tr><td align="center">Dynamic Matching</td><td align="center">Standard Matching</td></tr>
<tr><td>$R_\tau^0(t_s) = \{T_u(\infty) \geq t_s + \tau, T_s > t_s + \tau\}$</td><td>$R_\tau^0(t_s) = \{\min\left(T_u(\infty), T_u(T_s)\right) \geq t_s + \tau, T_s > t_s\}$</td></tr>
<tr><td>$R_\tau^1(t_s) = \{T_u(t_s) \geq t_s + \tau, T_s = t_s\}$</td><td>$R_\tau^1(t_s) = \{T_u(t_s) \geq t_s + \tau, T_s = t_s\}$</td></tr>
<tr><td>$X_\tau^0(t_s) = \{T_u(\infty) = t_s + \tau, T_s > t_s + \tau\}$</td><td>$X_\tau^0(t_s) = \{\min\left(T_u(\infty), T_u(T_s)\right) = t_s + \tau, T_s > t_s\}$</td></tr>
<tr><td>$X_\tau^1(t_s) = \{T_u(t_s) = t_s + \tau, T_s = t_s\}$</td><td>$X_\tau^1(t_s) = \{T_u(t_s) = t_s + \tau, T_s = t_s\}$</td></tr>
</table>

Looking at the two definitions of R^0 and X^0, one can see that the stan-

dard matching approach includes in the control group individuals who can enter treatment between t_s and $\tau + t_s$ (contamination effect). To measure the extend of this latter effect we estimate a contamination rate at each date t, defined as the ratio between the number of individuals that will be treated strictly after t and the number of individual still unemployed at t and not yet treated. Figure 2 shows this contamination rate as a function of t.

Figure 2: Contamination rate as a function of the treatment date
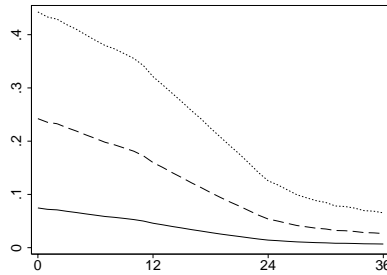
solid line: whole sample, dashed/dotted line: keeping 25%/10% of the non treated.

We see that contamination is small and declining over time. It starts at 8% and goes down to 5% after 12 months. In order to study the incidence of contamination on estimated survival functions, we also show in Figure 2 the results using artificial samples including all spells with treatment but only either 25% or 10% of treated spells without treatment. In both cases, the contamination is much stronger and remains decreasing with respect to the duration before treatment.

Figure 3: Bias of standard matching methods
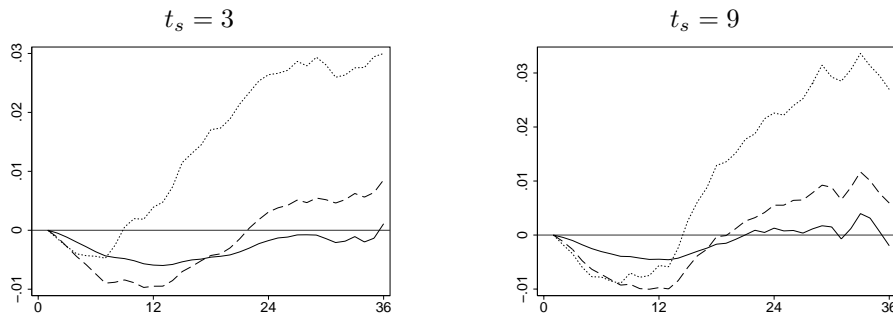
solid line: whole sample, dashed/dotted line: keeping 25%/10% of the non treated.

Figure 3 shows the bias of the estimated treatment effects $TT_G(t_s)$ for $t_s = 3$ and $t_s = 9$, and for three samples. We first see that the bias in the whole sample

10

is small. It is always negative and less than .5%. However when considering artificial samples with only 10% or 25% of the spells without treatment, the biases become more important. It would be even more important if the policy had a stronger effect.

4 Conclusion

We propose a methodological foundation for the use of matching techniques in the cases of dynamic assignment. We first emphasize the importance of the no anticipation assumption in defining a counterfactual and thus relevant treatment parameters. We then show that these parameters are identified using a typical conditional independence assumption on potential durations. We apply our method to training programs in France and detail the implementation of our estimate. We find that the contamination bias is small in our data. However, since a few individuals enter training, the contamination rate is itself small. When using artificial samples in which the contamination rates are higher we find substantial differences between our method and the standard matching approach.

In this paper, we only consider a single treatment and a single duration. It could be interesting to extend our estimators to multiple treatments and outcomes.

References

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

Abbring, Jaap and Gerard J. van den Berg (2004). "Analyzing the effect of dynamically assigned treatments using duration models, binary treatment models, and panel data models." *Empirical Economics*, 29, 5–20.

Carcillo, Stéphane and David Grubb (2006). "From inactivity to work: the role of active labour market policies." OECD Social Employment and Migration Working Papers, No. 36, OECD Publishing, Paris.

Cochran, William (1968). "The Effectiveness of Adjustments by Subclassification in Removing Bias in Observational Studies." *Biometrics*, 24(2), 295-313.

Crépon Bruno, Marc Ferracci, and Denis Fougère (2007). "Training the Unemployed in France: Impact on Unemployment Duration and Recurrence." Working paper, IZA.

Crépon, Bruno and Rosen Deplatz (2001). "Une Nouvelle Evaluation des Allégements de Charges Sociales sur les Bas Salaires." *Economie et Statistiques*, 348(8), 3-24.

Fredriksson, Peter and Per Johansson (2008). "Dynamic Treatment Assignment - The Consequences for Evaluations using Observational Data." *Journal of Business and Economic Statistics*, forthcoming.

Heckman, James J., Hidehiko Ichimura and Petra Todd (1998). "Matching as an econometric evaluation estimator." *Review of Economic Studies*, 65, 261–294.

Hirano, Keisuke and Guido Imbens (2004). "The Propensity Score with Continuous Treatments." In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, edited by A. Gelman and X.-L. Meng. New York: Wiley.

Hirano, Keisuke, Guido Imbens, and Geert Ridder (2003). "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica*, 71(4), 1161-89.

Imbens, Guido (2000). "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika*, 87(3), 706-10.

Lechner, Michael (2002). "Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies." *Review of Economics and Statistics*, 84(2), 205-220.

Lechner, Michael and Ruth Miquel (2005). "Identification of effects of dynamic treatments by sequential conditional independence assumptions." Working paper, University of St. Gallen.

Richardson, Katarina and Gerard J. van den Berg (2008). "Duration dependence versus unobserved heterogeneity in treatment effects: Swedish labor market training and the transition rate to employment." Working paper, IFAU, Uppsala.

Rosenbaum, Paul and Donald Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 70(1), 41-55.

Rubin, Donald (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology*, 66(5), 688-701.

Sianesi, Barbara (2004). "An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s." *Review of Economics and Statistics*, 86(1), 133-55.