DISCUSSION PAPER SERIES

IZA DP No. 3848

Active Labor Market Policy Effects in a Dynamic Setting

Bruno Crépon
Marc Ferracci
Grégory Jolivet
Gerard J. van den Berg

November 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Active Labor Market Policy Effects in a Dynamic Setting

Bruno Crépon
CREST-INSEE, CEPR and IZA

Marc Ferracci
University of Marne-la-Vallée and CREST-INSEE

Grégory Jolivet
University of Bristol

Gerard J. van den Berg
VU University Amsterdam, IFAU-Uppsala, Netspar, IFS, CEPR and IZA

Discussion Paper No. 3848
November 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3848
November 2008

ABSTRACT

Active Labor Market Policy Effects in a Dynamic Setting*

This paper implements a method to identify and estimate treatment effects in a dynamic setting where treatments may occur at any point in time. By relating the standard matching approach to the timing-of-events approach, it demonstrates that effects of the treatment on the treated at a given date can be identified even though 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 until treatment and the counterfactual durations until 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.

JEL Classification: J64, C21, C31, C41, C14

Keywords: treatment, program participation, unemployment duration, training, propensity score, matching, contamination bias

Corresponding author:

Gerard J. van den Berg
Department of Economics
VU University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
E-mail: gjvdberg@xs4all.nl

* We thank Marco Caliendo, Per Johansson, Johan Vikström, and participants in the EEA 2008 Invited Session “Recent Developments in Program Evaluation" for useful comments. Part of this project was carried out when Jolivet was at the VU University Amsterdam as postdoctoral fellow of the EU-RTN “Microdata" project.

1 Introduction

Active labor market policy (ALMP) programs for the unemployed include job search assistance, training, and subsidized work (see Carcillo and Grubb, 2006, for a recent overview). A typical feature of ALMP is that 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 formal 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 as a methodological implication that 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 for those who are not yet treated at t_s. Application of standard evaluation methods like matching 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 reatment effects. In particular, under the assumption of unconfoundedness (CIA) of potential outcomes and assigned treatment conditional on covariates, matching is 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.

Fredriksson and Johansson (2008) deal with this by developing a matching estimator for average effects of treatment at t_s on the remaining 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. See also De Luna and Johansson (2007)'s estimator.

In this paper we clarify and advance on the literature. We develop an evaluation framework with counterfactual duration outcomes à la Abbring and Van den Berg (2003). 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 can be estimated with matching, given unconfoundedness and the no-anticipation assumption. The estimator is similar to the Fredriksson and Johansson (2008) estimator, but we also relate it to Abbring and Van den Berg (2003)'s Timing-of-Events framework for dynamic treatment assignment and duration outcomes. 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. Also, alternative methods are available. One may estimate semi-parametric models for outcomes and treatments at various points in time, to obtain estimates of the treatment effect as a function of the time elapsed since enrollment, the elapsed unemployment duration, and unobserved-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).

2 Identification of dynamic treatment effects

2.1 The model

Consider individuals who enter a given state U (say unemployment) at date 0. 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 effects of receiving the treatment at date t_s.

We therefore consider the *potential* durations $T_u(t_s)$ spent in U when treatment occurs at the assigned 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 take time as discrete.

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 proceed, we adopt the "no anticipation" assumption from Abbring and Van den Berg (2003)'s 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 given individual, all counterfactual processes have the same distribution up to the first treatment date among them. In particular,

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

$T_u(\infty)$ can be viewed as the duration if the individual's treatment will always be "later". This duration is the counterfactual corresponding to "no treatment".

By analogy to the matching literature, we aim to identify and estimate the average effect ATT of treatment at t_s on the treated, with re-employment between t_s and $\tau + t_s$ as the outcome of interest,

$$TT_{G_\tau}(t_s) = E\left[G_\tau\left(T_u(t_s)\right) - G_\tau\left(T_u(\infty)\right) \middle| T_s = t_s, T_u(t_s) > t_s\right]. \tag{2}$$

with $G_\tau(t) = 1\{t > \tau + t_s\}$, for any $\tau > 0$.

This deviates in a number of ways from the Timing-of-Events approach. In the latter, the interest is in (average) treatment effects on the counterfactual distributions, which are defined regardless of the actual assignment mechanism. Secondly, in the latter approach, the primary interest is in effects on individual hazard rates rather than survival probabilities. Thirdly, the Timing-of-Events approach provides a comprehensive set of estimates for effects at all t_s, at the expense of semi-parametric assumptions, whereas in the matching approach the ATT estimates at different t_s concern sub-populations that are systematically different from each other. This is due to the dynamic selection driven by unobserved heterogeneity in treatment effects, and applies even in the case of randomized assignment.[2]

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

[2] Note that, since TT involves survival functions, its sign leads to a reverse interpretation than usual. If TT is positive, treatment at date t_s decreases the probability of leaving before $t_s + \tau$. If the treatment concerns training, then a positive TT parameter indicates that training tends to lengthen unemployment.

2.2 Identification

The TT expression involves two quantities. The first one is the average outcome for those who are observed to be treated at t_s:

$$E\left[G_\tau\left(T_u(t_s)\right)\middle|T_s = t_s, T_u(t_s) > t_s\right] \tag{3}$$

This is identified from outcomes of those observed to be treated at t_s. If durations can be right-censored then we need an additional assumption (see below).

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

$$E\left[G_\tau\left(T_u(\infty)\right)\middle|T_s = t_s, T_u(t_s) > t_s\right]. \tag{4}$$

Its identification gives rise to 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 conditional on observed individual characteristics X, assignment to treatment is independent of the counterfactual. Specifically, in line with Abbring and Van den Berg (2003), unconfoundedness is assumed through conditional independence of the *latent* variable T_s from the joint counterfactuals $\{T_u(t_s)\}$,

$$\{T_u(t_s); t_s \geq 0\} \quad \perp \quad T_s \quad | \quad X. \tag{A2}$$

The two assumptions (A1) and (A2) allow us to identify the missing term (4), in two steps. First, assumption (A2) suggests to consider individuals who are at risk and not yet treated at t_s as a potential control group:

$$E\left[G_\tau\left(T_u(\infty)\right)\middle|T_s = t_s, T_u(t_s) > t_s, X\right] = E\left[G_\tau\left(T_u(\infty)\right)\middle|T_s > t_s, T_u(t_s) > t_s, X\right].$$

Then, for individuals in the control group, we effectively observe $T_u(\infty)$ if individuals leave U before entering treatment, or we observe $T_u(\infty)$ 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 $T_u(\infty)$. This information is sufficient to identify the distribution of $T_u(\infty)$ and therefore the ATT.

The duration in state U may be right-censored. For example, if U is unemployment, censoring occurs if the spells exceed the date up to which information has been collected. To deal with this issue, we make another conditional independence assumption. With T_C denoting 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 this assumption, the distribution of $T_u(t_s)$ conditional on being treated at date t_s and X is still identified. Therefore, the two terms of TT are both identified as well. Note that Assumption A3 is stronger than the usual assumption on right-censoring in duration analysis, which does not require independence but merely uninformativeness.

2.3 Estimation

We present a simple two-stage method to estimate our parameters of interest. In policy evaluations with standard matching, it is well known that the dimensionality of X can be reduced by the propensity score property (cf. Rosenbaum and Rubin, 1983). This can be extended to the case of multi-valued treatments, as follows:[3] assuming that there is an index $s(X)$ such that $f_{T_s|X}(t) = f(t, s(X))$. Then the unconfoundedness assumption (A2) implies independence conditional on $s(X)$ as well. In the first estimation step we estimate the propensity score. Note that the duration up to treatment is observed for individuals entering treatment and is censored for individuals leaving unemployment before treatment.

Next, we proceed to the matching step. 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. We match individuals on the score $s(X)$. Initial populations of treated and non-treated can be split into subpopulations with similar values of the score.[4] We depart from the matching literature (e.g. Sianesi, 2004) and use blocking methods (see Cochran, 1968, or Rosenbaum and Rubin, 1983). This choice is driven by practical considerations (it allows for computationally fast estimation) and seems reasonable in the light of the large sample sizes we have. 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.[5] 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 obtain the survival function as

[3] See Crépon and Desplatz (2003) and Hirano and Imbens (2004).

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

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

$(1-h_{t_s}) \times (1-h_{t_s+1}) \times \cdots \times (1-h_{t_s+\tau})$. The difference between the survival functions of treated and non-treated is averaged using the distribution of the score function in the treatment population.

3 An empirical illustration concerning 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.

We include a rich set of covariates in the X vector to ensure that the unconfoundedness assumption (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 Section 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 $t_s = 9$.

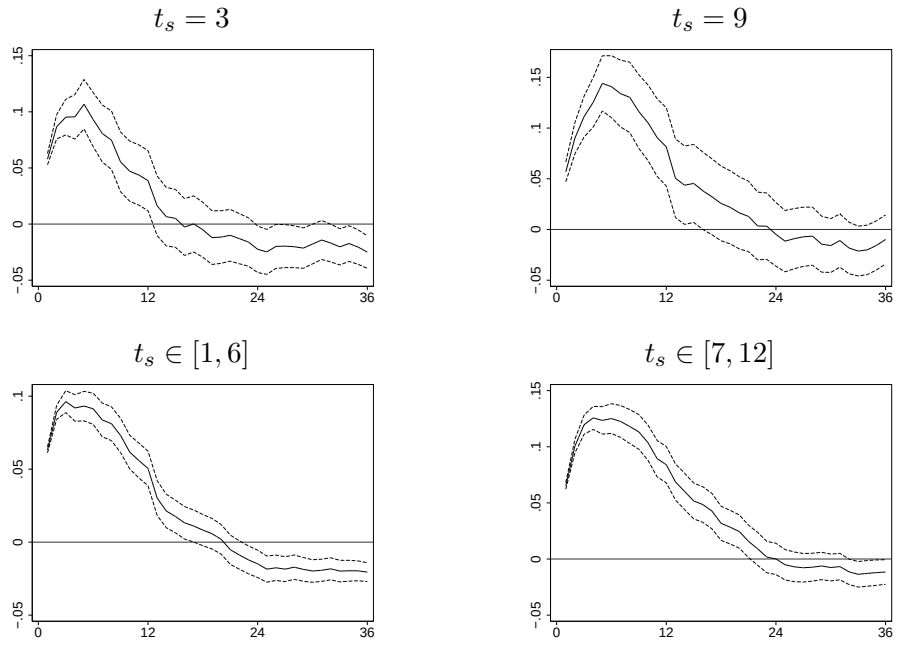
The figures clearly show that the effect at these two treatment dates have the same pattern. There is first a locking in period that lasts around 18 months. The effect is important as the survival rate can increase by 10%. After this locking in period, the effect of training on survival is negative. 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 was already pointed out in numerous studies: training the unemployed does substantially shorten their unemployment spells.

We also estimate more aggregated parameters such as:

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

for treatment dates ranging from $t_{s1} = 1$ to $t_{s2} = 6$ and from $t_{s1} = 7$ to $t_{s2} = 12$. 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



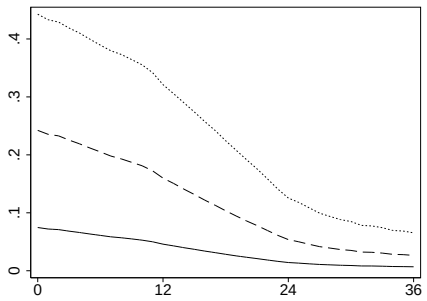
There are two main differences between our dynamic matching method and the standard matching approach. The first one is related to exits: in our setting we only count exits without censoring, while the standard matching method also includes them. More importantly, the second difference lies in the definition of the risk sets (and thus of exits). To fix ideas, let $R^S_\tau(t_s)$ be the risk set at date $\tau + t_s$ for individuals with treatment status $S = 0, 1$, for the treatment date t_s. Let $X^S_\tau(t_s)$ 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:

Dynamic Matching	*Standard Matching*
$R^0_\tau(t_s) = \{T_u(\infty) \geq t_s + \tau, T_s > t_s + \tau\}$	$R^0_\tau(t_s) = \{\min(T_u(\infty), T_u(T_s)) \geq t_s + \tau, T_s > t_s\}$
$R^1_\tau(t_s) = \{T_u(t_s) \geq t_s + \tau, T_s = t_s\}$	$R^1_\tau(t_s) = \{T_u(t_s) \geq t_s + \tau, T_s = t_s\}$
$X^0_\tau(t_s) = \{T_u(\infty) = t_s + \tau, T_s > t_s + \tau\}$	$X^0_\tau(t_s) = \{\min(T_u(\infty), T_u(T_s)) = t_s + \tau, T_s > t_s\}$
$X^1_\tau(t_s) = \{T_u(t_s) = t_s + \tau, T_s = t_s\}$	$X^1_\tau(t_s) = \{T_u(t_s) = t_s + \tau, T_s = t_s\}$

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

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

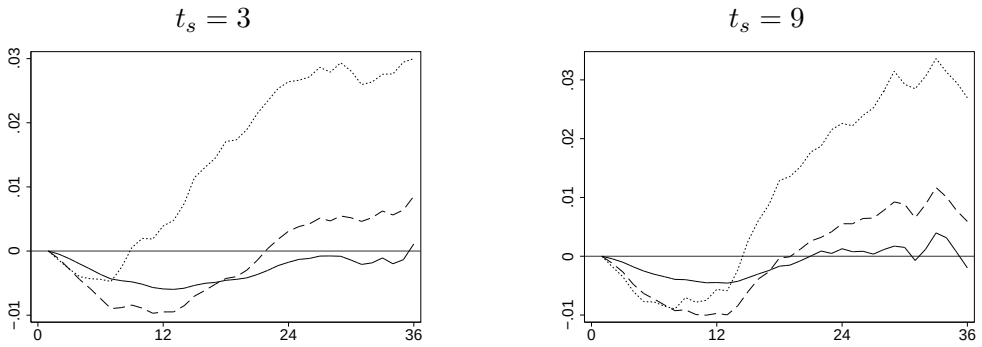


solid line: whole sample, dashed/dotted line: keeping 25%/10% of the non treated.

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 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 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.

Figure 3: Bias of standard matching methods



solid line: whole sample, dashed/dotted line: keeping 25%/10% of the non treated.

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 one treatment and one 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.

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, 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 Rozenn Desplatz (2003). "The Effects of Payroll Tax Subsidies for Low Wage Workers on Firms Level Decisions." Working Paper, CREST.

De Luna, Xavier and Per Johansson (2007). "Matching estimators for the effect of a treatment on survival times." Working paper, IFAU.

Fredriksson, Peter and Per Johansson (2008). "Dynamic Treatment Assignment - The Consequences for Evaluations using Observational Data." *Journal of Business and Economic Statistics*, 26, 435–445.

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, 1161–1189.

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.

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

Rubin, Donald (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology*, 66, 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), 133–155.