published in The Journal of the Royal Statistical Society
Series A, 173, 67-92, 2010

UNEMPLOYED AND THEIR CASEWORKERS:

SHOULD THEY BE FRIENDS OR FOES?

Stefanie Behncke*, Markus Frölich+, and Michael Lechner*

* Swiss Institute for Empirical Economic Research (SEW), University of St. Gallen

+ University of Mannheim and IZA and SEW

Abstract

In many countries, caseworkers in public employment offices have dual roles of counselling and monitoring unemployed persons. These roles often conflict, which results in important caseworker heterogeneity: Some consider providing services to their clients and satisfying their demands as their primary task. However, others may pursue their own strategies, even against the will of the unemployed person. They may assign jobs and labour market programmes without the consent of the unemployed person. Based on a very detailed *linked jobseeker-caseworker* dataset for Switzerland, we investigate the effects of caseworkers' cooperativeness on the employment probabilities of their clients. Modified statistical matching methods reveal that caseworkers who place less emphasis on a cooperative and harmonic relationship with their clients increase their employment chances in the short and medium term.

Keywords: Public employment services, unemployment, statistical matching methods

JEL classification: J68, C31

Stefanie Behncke, Michael Lechner
Swiss Institute for Empirical Economic Research (SEW)
University of St. Gallen,
Varnbüelstr. 14, CH-9000 St. Gallen, Switzerland
Stefanie.Behncke@unisg.ch, Michael.Lechner@unisg.ch
www.sew.unisg.ch/lechner

Markus Frölich
Universität Mannheim
L7, 3-5
68131 Mannheim, Germany
froelich@uni-mannheim.de
froelich.vwl.uni-mannheim.de

1 Introduction

In most countries, caseworkers in unemployment offices have dual roles of *counselling* and *monitoring* unemployed persons. These two roles often conflict with each other. Caseworkers must establish a trustful and empathetic relationship with their clients for providing effective counselling. However, they must also police job search behaviour and initiate and enforce sanctions if it falls short of certain requirements. Some caseworkers pursue a more dominating and demanding stance vis-à-vis the unemployed, while others aim at a more cooperative relationship, devoid of conflicts. Depending on their attitude toward cooperation, caseworkers might differ in their assignment of jobs, active labour market programmes, imposition of sanctions, in addition to more personal channels such as counselling style. This in turn is likely to affect the employment prospects of their clients.

In this paper, we examine how different degrees of caseworker's cooperativeness affect the employment chances of their clients by exploiting a unique and very informative *linked jobseeker-caseworker* dataset for Switzerland. The econometric analysis is based on matching estimators.

Thus far, the effects of single instruments on employment were considered separately: positive effects were found for *sanctions* (Van den Berg et al., 2004; Lalive et al., 2005; Abbring et al., 2005) and for the assignment of onerous labour market programmes via *threat effects* (Black et al., 2003; Graversen and van Ours, 2006; Lalive et al., 2006; Lalive and Zehnder, 2007), no or positive effects were found for *monitoring* (Ashenfelter et al., 2000; Bloom et al., 2003; Meyer, 1995; Gorter and Kalb, 1996; Dolton and O'Neill, 1996), mixed effects were found for active *labour market programmes* (Heckman et al., 1999; Martin and Grubb, 2001; Wunsch, 2005). We augment this literature by analysing the relationship between caseworker and unemployed as a whole. A more demanding caseworker may use certain instruments more often, but the different counselling styles by themselves can have important effects. Thus, in addition to labour market policy instruments, the personal relationship between the caseworker and the unemployed person may have an important impact on motivation, job search intensity, and job acceptance. We contribute to the traditional evaluation literature by considering the imposition of sanctions or the assignment of labour market programmes as an outcome of the caseworker's attitude towards cooperation.

Our estimation results indicate a positive effect of reduced caseworker cooperativeness on the average employment probabilities of their clients of about 2 percentage points. Hence, pursuing a more demanding stance vis-à-vis unemployed persons increases employment probabilities by a non-negligible amount. Positive effects are found in the short and medium term up to 3 years after becoming unemployed. This increased employment is not obtained at the cost of reduced stability of jobs. The sensitivity of these results is explored by examining several alternative specifications, in particular concerning the choice of the control variables and various definitions of the treatment variable. The results are rather stable, although in several cases precision becomes an issue.

The structure of the paper is as follows: The following section discusses the unemployment insurance system in Switzerland. Section 3 describes the data, and Section 4 the statistical methodology. Sections 5 and 6 give the main estimation results, and Section 7 concludes. Several appendices provide additional details.

2 The Swiss labour market and the role of caseworkers

2.1 Overview

Until the recession of the early 1990s, unemployment was extremely low in Switzerland, a small country with 26 different administrative regions, called *cantons*. As shown in Figure 1, with the recession, the unemployment rate rose rapidly to 5%. This triggered a comprehensive revision of the Federal Unemployment Insurance Act in 1996. The municipal employment offices, which had largely been concerned with administering unemployment benefits, were consolidated to r*egional employment offices* with the aim of providing professional services with respect to counselling, placement, activation, and training. Caseworkers were hired and especially trained for these purposes.

Figure 1: Unemployment rate in Switzerland from January 1990 to November 2008

Note: Monthly unemployment rate, January 1990 – November 2008, Source: Swiss National Bank Monatshefte.

2.2 Caseworkers' incentives and autonomy

The federal State Secretariat for Economic Affairs (SECO) sets clear targets regarding the goals, which all employment offices and caseworkers should pursue. Every year, a performance indicator is computed for every employment office as a weighted average of four indices: Rapid re-employment (50% weight), avoidance of long-term unemployment (20%), avoidance of benefit exhaustion (20%), and avoidance of repeated unemployment (10%). The computed performance, based on this indicator, has no financial consequences for the regional employment office, but affects its reputation. However, neither the federal nor the cantonal governments provide strict guidelines as to how to reach these targets. Regional employment offices enjoy a substantial autonomy in the implementation of the unemployment insurance law. Moreover, caseworkers generally have considerable leeway when dealing with their clients. Many employment office managers consider it important that their caseworkers develop their own counselling style and react to the needs of their clients without being bound by many bureaucratic rules (Frölich et al., 2007). While in some employment offices the performances of individual caseworkers' are measured by this indicator as well, there are no performance-based pay components. Yet, caseworkers still have an incentive to perform well, as they face a strong "business cycle": when the national unemployment rate drops from 5.5% to only 1.5% within four years (see Figure 1), caseworkers are in excess supply. Without any strict guidelines, however, they strive to reach their targets according to their own beliefs and preferences.

2.3 The relationship between caseworker and unemployed persons

The relationship between the caseworker and his client is characterised by the two roles of the caseworker: to help the unemployed person in searching and finding appropriate employment and to monitor whether the unemployed person searches thoroughly enough and is indeed willing to take up any job offer with acceptable pay and within acceptable commuting distance. Some caseworkers put more emphasis on their role as a counsellor and aim for a trustful relationship, whereas other caseworkers see their policing role as being more important and are more dominating and demanding vis-à-vis the unemployed person.

To analyse the effects of the caseworker-client relationship a written questionnaire was administered to all caseworkers in Switzerland (for details, see Frölich et al. (2007)). A key question, shown in Table 1, asked the caseworker how important he/she considers cooperation with the client:

Table 1: Survey question on cooperativeness of the caseworker

How important do you consider the cooperation with the jobseeker, regarding placements in jobs, and assignment of active labour market programmes?
☐ 1 Cooperation is very important; the wishes of the unemployed person should be satisfied.
☐ 2 Cooperation is important, but placements in jobs and active labour market programmes should sometimes be assigned or declined in spite of the unemployed person's wishes.
☐ 3 Cooperation is less important; I should assign placements in jobs and active labour market programmes independent of the wishes of the unemployed person

Note: English translation. Questionnaires were in German, French, and Italian.

52% of the caseworkers chose option one, 39% of caseworkers chose option two, and 9% of caseworkers chose option three. Only very few caseworkers did not respond to this particular question. They are dropped from the analysis.

When comparing these answers with the responses to other items of the questionnaire we observe that the less cooperative caseworkers consider control and sanctions, job assignments, and employment programmes as instruments that are more important, while counselling meetings and temporary wage subsidies are considered less important. They also responded that they tended to assign active labour market programmes to apply pressure and to control their clients' availability for jobs.

The cooperation attitude is likely to contain some measurement error with respect to the actual cooperativeness towards a particular client. First, cooperativeness is self-reported by the caseworker. Second, the cooperativeness of a caseworker may also vary with the characteristics of the client. However, the survey question did not differentiate by client type and thereby only enquired the average cooperativeness. We generally expect that measurement error is uncorrelated with the true cooperation type, as we believe that there is no systematic misreporting depending on the true type. This would result in the so called *attenuation bias* which biases the estimated effect towards zero.

3 Data

3.1 Data and sample selection

The population consists of all individuals who registered at Swiss regional employment offices anytime during the year 2003. For them, very detailed individual information is available from the databases of the unemployment insurance system and social security records. They contain socio-economic characteristics including nationality, qualification, education, language skills, experience, profession, position, and industry of last job, occupation and industry of desired job, an employability rating by the caseworker, etc. The data also contain information on registration and de-registration of unemployment, benefit payments and sanctions, participation in active labour market programmes (ALMP), and employment histories since January 1990 with monthly information on earnings and employment status.

In total, 239,004 persons registered as new jobseekers during 2003. For each person we consider only the first registration in 2003. Any further registrations are considered as outcomes of the caseworker contacts following the first registration. We exclude jobseekers who have not claimed benefits and individuals who applied for or claim disability insurance, as these groups are largely immune to potential caseworker threats and sanctions. Furthermore, we exclude foreigners without permanent or yearly work permit, as they are not entitled to most of the services of unemployment insurance. We exclude unemployed whose caseworkers are undefined, which may happen if an unemployed person de-registers before being assigned to a caseworker. We also exclude a few employment offices that are not comparable to other offices. In our main analysis we focus on the prime-age group (24 to 55 years old), with a final sample size of 100,222 unemployed persons. See Appendix A for further details.

3.2 Definition of outcomes and treatment variables

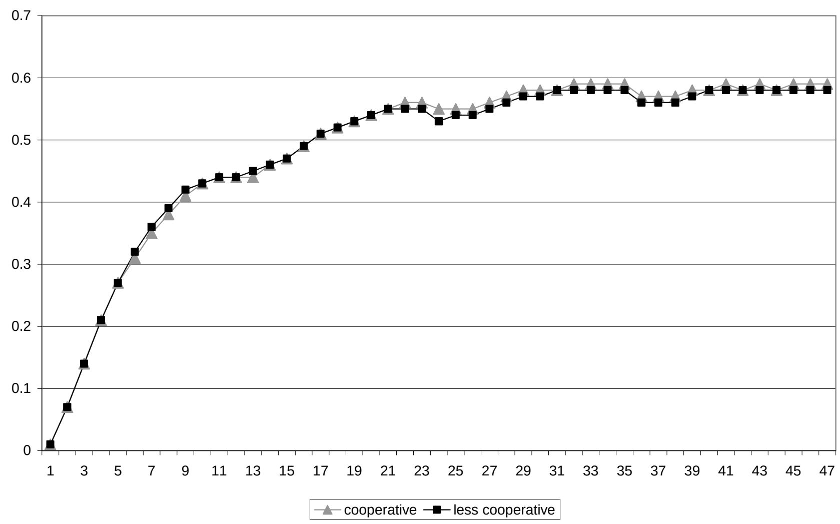
Each newly registered unemployed person in 2003 was linked to his first caseworker. Usually the same caseworker remains in charge for the entire spell of unemployment. Still, we focus on the first caseworker to avoid any concerns about endogenous caseworker changes. Information on the caseworker stems from a written questionnaire, which was sent to all caseworkers who were employed at an employment office in 2003 and were still active in December 2004, i.e. at the time the questionnaire was sent. The questionnaire contained questions about aims, strategies, processes, and organisation of the employment office and the caseworkers. 1,560 caseworkers and managers of employment office returned the questionnaire, yielding a response rate of 84%. Of the 1560 individuals who returned the questionnaire, 159 office managers who did not counsel jobseekers during 2003 were not asked the key question shown in Table 1. A further 16 caseworkers did not answer this question.

In our main specification, we consider those who answered with option 1 as *cooperative* and those who chose option 2 or 3 as less cooperative caseworkers. Although the cooperation attitude of a caseworker may vary between his clients, we expect that a cooperative caseworker be more cooperative to all his clients than a less cooperative caseworker is. Combining options 2 and 3 to define non-cooperation is based on the presumption that both show a deviation from the full cooperation attitude that more than half of the caseworkers display. In our robustness analysis, we will also consider other definitions; in particular, we compare option 1 directly with option 3.

An individual is considered as *employed* in month t if she has de-registered at the employment office *and* if the exit state is known to be employment. The employment status, $Y_{i,t_0+\tau}$, is measured monthly until the end of 2006, relative to the time of first registration, t_0. Since for each individual t_o is some month in 2003, we observe then his employment status for at least three years. This allows for an estimation of not only the short-term effects, but also the estimation of medium-term effects of co-operation.

Figure 2 shows the employment rate for our sample relative to the time of registration at the employment office. The black line represents the employment rates for those unemployed who were counselled by a less cooperative caseworker, whereas the grey line refers to those counselled by cooperative caseworkers. The employment rates for both groups of unemployed are very similar. About 1% are employed again one month after registration and about 44% after one year.

Figure 2: Average employment rate in month t after registering as unemployed



Note: Average employment rates are for prime age individuals in main sample. The grey (black) line shows averages for the 51923 (48299) unemployed who were counselled by a cooperative (less cooperative) caseworker. The first 36 months are observed for the full sample. The months 37 to 47 are observed only for subsamples of decreasing size. Month 37 is observed only for those individuals who registered as unemployed in the period January to November 2003. Month 47 is observed only for those who registered unemployed in January 2003.

The employment rates in Figure 2, however, do not account for the fact that the unemployed counselled by cooperative and less cooperative caseworkers are quite distinct. Table 2 shows that these two groups of unemployed differ in many characteristics. Amongst those unemployed whose caseworkers do not regard counselling to be important there is a relatively higher incidence of foreigners, unskilled, those whose employability rating is low and those who live in regions with higher local unemployment rate. In this sense the more difficult cases in terms of employment opportunities are more likely to go to less co-operative case workers than other cases. If one would control for these differences, then it is likely that the average employment rate of the group with less cooperative caseworkers in Figure 2 would move upwards. We also observe clear differences by language region: amongst those unemployed who are confronted with less cooperative caseworkers relatively more are from the Italian-speaking and relatively fewer from the German and French speaking regions. The language region will therefore be an important control variable, and we will use interaction terms with language regions in our later regressions throughout.

We are also able to compare the allocation process between unemployed whose caseworkers are from different types, as we know from the questionnaire how unemployed were assigned to caseworkers. The answers "at random", "alphabetically" and "by caseload" (i.e. number of clients) essentially mean that allocation was *not* based on unobserved characteristics of the jobseeker. The answers "by industry sector", "by occupation group", "by age of the unemployed", "by employability", "by region", "other" indicate that allocation was based at least partly on observed (and perhaps unobserved) characteristics of the jobseeker. We observe that amongst those unemployed whose caseworkers are non-cooperative there is a relatively higher incidence that caseworkers are specialised towards counselling unemployed of a certain industry or occupational group.

Lastly as Table 3 shows, caseworkers also differ in their own characteristics, which may be related to their efficacy in counselling and placing unemployed. The non-cooperative caseworkers are on average younger and have participated in the vocational training programme for caseworkers more often.

Table 2: Selected characteristics of the unemployed by cooperation attitude of the caseworker

	Caseworker considers cooperation as		
	Very important (option 1)	Important, but (option 2)	Less important (option 3)
Characteristics of the unemployed person			
N (number of unemployed)	51,923	39,310	8,989
Female	0.45	0.44	0.42
Age	36.5	36.6	36.6
Swiss	0.63	** 0.61	*** 0.56
Foreigner with permanent work permit	0.24	* 0.25	*** 0.29
Foreigner with yearly work permit	0.13	* 0.14	*** 0.16
Qualification: unskilled	0.21	0.23	*** 0.28
Qualification: semiskilled	0.16	0.16	* 0.14
Qualification: skilled without degree	0.04	0.04	0.05
Qualification: skilled with degree	0.58	0.57	*** 0.53
Employability rating: low	0.13	0.14	* 0.17
Employability rating: medium	0.75	0.74	0.72
Employability rating: high	0.12	0.12	0.11
Local labour market characteristics			
German speaking employment office	0.69	0.71	0.60
French speaking employment office	0.26	0.22	0.19
Italian speaking employment office	0.05	0.07	*** 0.20
Cantonal unemployment rate	3.70	3.75	3.79
Unemployment rate in industry	4.83	* 4.93	*** 5.11
Allocation of unemployed to caseworkers [a)]			
At random	0.22	0.23	0.22
By alphabet	0.03	0.04	0.07
By number of clients	0.42	0.44	0.44
By industry	0.52	0.57	0.56
By occupation	0.52	*** 0.62	0.57
By age	0.03	0.04	0.01
By employability	0.07	0.06	0.07
By region	0.12	0.11	0.17
Other	0.08	0.08	0.04

Note: The entries in the table are proportions for categorised and binary variables, means for interval level variables or number of observations, by subgroup of cooperation attitude. ***, **, * indicate significant differences between column 2 versus column 1 and between column 3 versus column 1, respectively, at the 1%, 5% and 10% level, respectively. Significance tests are based on robust estimates of standard errors that take into account that unemployed counselled by the same caseworker are not independent observations (see Rogers,1993).
[a)] Caseworkers were asked according to which criteria unemployed were assigned to them. Multiple answers to this question were permitted. Hence, the means do not sum up to 1.

Table 3 Selected characteristics of the caseworkers by their cooperation attitude

	Caseworker considers cooperation as		
	Very important (option 1)	Important, but (option 2)	Less important (option 3)
Characteristics of the caseworker			
N (number of caseworkers)	723	540	122
Female	0.44	0.42	0.41
Age	46.5	*** 44.4	*** 44.1
Tenure in employment office in years	5.93	6.21	5.88
Own experience of unemployment	0.65	0.61	0.65
Education: vocational training	0.26	* 0.30	** 0.35
Education: above vocational training	0.45	0.43	0.44
Education: tertiary track (university or polytechnic)	0.29	0.27	* 0.21
Degree in vocational training for caseworkers	0.19	** 0.23	0.24

Note: The entries in the table are proportions for categorised and binary variables, means for interval level variables or number of observations, by subgroup of cooperation attitude. ***, **, * indicate significant differences between column 2 versus column 1 and between column 3 versus column 1, respectively, at the 1%, 5% and 10% level, respectively.

4 Methodology

4.1 Conditional independence assumption as identification strategy

Consider an individual i who registers as unemployed at time t_0 at the nearest regional employment office. This person is then assigned to a caseworker of that office. The caseworker is of a particular type with respect to his willingness to cooperate with his client. Let D_i denote the attitude of the caseworker who is counselling individual i. $D_i = 1$ represents a non- or less-cooperative caseworker whereas $D_i = 0$ represents a cooperative caseworker.

We are interested in the impact of a cooperative caseworker on the subsequent employment prospects of this unemployed person, which we measure by the employment status, $Y_{i,t_0+\tau}$, in month τ after registration. In particular, we would like to compare the employment status with the potential employment status if the same unemployed person was counselled by a caseworker with a different attitude. We base our analysis on the prototypical model of the statistical evaluation literature with a binary treatment variable D (see Neyman, 1923, Fisher, 1935, Rubin, 1974, 1979). Let

$$Y^d_{i,t_0+\tau} \tag{1}$$

be the potential outcome τ months after unemployment registration at month t_0, if the caseworker was of type d. In other words, $Y^0_{i,t_0+\tau}$ is the employment outcome that would have been observed had person i been counselled by a cooperative caseworker, whereas $Y^1_{i,t_0+\tau}$ is the employment outcome that would have been observed had person i been counselled by a non-cooperative caseworker. Since outcomes are always considered relative to time of registration t_0 which is treated as an additional covariate of person i, we neglect the time subscripts without loss of generality. We will further simplify notation by dropping the individual person subscript.

The average treatment effect for a person who has been counselled by a non-cooperative or by a cooperative caseworker is

$$E[Y^1 - Y^0 \mid D = 1] \qquad \text{(ATET)},$$
$$E[Y^1 - Y^0 \mid D = 0] \qquad \text{(ATEN)}.$$

We will refer to these parameters as the average treatment effect on the treated (ATET) and the average treatment effect on the non-treated (ATEN), respectively. The following discussion focuses on the ATET, with obvious modifications for the ATEN.

In order to be able to estimate the expected potential outcomes for different values of d, we need to observe variation in D that is exogenous with respect to the outcome variable. The observed type of the caseworker D might be related to many factors that also have an impact on employment chances, such that in general

$$E[Y^d] \neq E[Y^d \mid D = d]. \tag{2}$$

However, if we were to condition on all variables *X* that determined the type of the caseworker and the potential employment chances of the unemployed person, the potential outcomes conditional on *X* would be identified as:

$$E[Y^d \mid X = x] = E[Y^d \mid X = x, D = d] \qquad \forall x \in \chi, \tag{3}$$

where $\chi \subseteq Supp(X)$ is a subset of the support of the random variable *X*. This assumption is referred to as the conditional independence assumption (CIA). We assume the CIA to hold for every value of *x* that lies in the support of *X* in the *D=1* and the *D=0* population, i.e. $\chi = Supp(X \mid D = 1) \cap Supp(X \mid D = 0)$.

The most crucial aspect of the identification strategy thus relies on being able to observe all confounding variables *X*. To do so, the very detailed linked caseworker-client dataset, described above, is essential together with an understanding of the determinants of the cooperation attitude of the caseworker. The cooperativeness *D* of the caseworker depends on three processes: First, which types of caseworkers are hired, second, how caseworkers are allocated to the unemployed, and third how their attitudes develop after having been trained and gaining experience on the job. In that caseworkers' attitudes may be related to their general skills of finding jobs for their clients, we include caseworker characteristics such as their age, gender, education, work experience, and experience of own unemployment as covariates. We also control for the specific variables of the unemployed allocation process to caseworkers, which we know from the questionnaire. A further aspect is that caseworkers not only differ in personality, but they also react to types of unemployed and the labour market environment they face. If vacancies are scarce and rapid re-employment appears difficult, caseworkers may be less demanding than in a more favourable environment. Similarly, a caseworker who counsels mainly individuals with a low employability rating may react differently than a caseworker, who for example is responsible mainly for youth. Therefore, we include in the analysis a large number of covariates on the unemployed person's employment history and the local labour market.

4.2 Nonparametric matching estimation

The estimator used is a matching estimator as implemented in Lechner et al. (2006). The advantage of matching estimators is that they are essentially nonparametric and that they allow for arbitrary individual effect heterogeneity. The estimator is nonparametric in the sense that the conditional expectation function E[$Y|X,D$] is not parametrically specified, i.e. the conditional expectation function is not assumed to belong to a class of functions that can be characterised by a finite dimensional parameter vector. In addition, the *X* do not have to be exogenous (Frölich 2008).

We first discuss identification of the ATET, before outlining the approach to estimating it. As argued before, simply comparing the average outcomes of *D*=1 and *D*=0 subsamples does not take into account that these individuals may differ not only in the cooperation attitude of their caseworker but also in other characteristics. The ATET can be written as:

$$
\begin{aligned}
E[Y^1 - Y^0 \mid D=1] &= E[Y^1 \mid D=1] - E[Y^0 \mid D=1] \\
&= E[Y \mid D=1] - E[Y^0 \mid D=1].
\end{aligned}
$$

Estimating the first term from the sample mean of *Y* from observations with *D=1* is straightforward. With $F_{X|D=1}$ as the conditional distribution function of *X* given *D=1*, we can write the second counterfactual term as $\int E[Y^0 \mid X, D=1] dF_{X|D=1}$. Due to the conditional independence assumption (3) we can replace *D=1* with *D=0* in the expectation here and hence Y^0 by *Y*. The ATET can now be written $E[Y \mid D=1] - \int E[Y \mid X, D=0] dF_{X|D=1}$. This expression for ATET depends now only on distributions of observable Y, D and X and it is thus identified. Hence, if we had access to an infinitely large sample we could estimate the ATET without error.

We consider now the estimation of $E[Y \mid X, D=0]$ written more fully as $E[Y \mid X=x, D=0]$ and which we will denote by *$m_0(x)$*. If the number of variables in *X* was small and they were categorised with few distinct values this could be achieved by mean values of *Y* within a few subgroups. With many variables in *X*, estimation by subsample averages is either not feasible anymore (there might be zero observations for some values of *x*) or probably very noisy. Some form of smoothing over neighbouring values of *x* is needed. One could impose some parametric functional form restriction, e.g. that $E[Y \mid X=x, D=0]$ is linear in *x*, and estimate a linear regression. A more flexible approach, if one does not know the true form of $E[Y \mid X=x, D=0]$, is a *nonparametric* regression of *Y* on *X* using only observations with *D*=0 and with covariate values in a neighbourhood of *x*. We return to the formulation of an estimator $\hat{m}_0(x)$ below. Given this estimator and estimating $F_{X|D=1}$ by the empirical distribution function over the N_1 cases in the sample with *D=1* we can now derive an estimator of ATET as the mean of *Y*- $\hat{m}_0(x)$ over cases with *D=1*:

$$\widehat{E[Y^1 - Y^0 \mid D=1]} = \frac{1}{N_1} \sum_{i:D_i=1} (Y_i - \hat{m}_0(x_i)) . \tag{4}$$

We turn now to the estimator $\hat{m}_0(x)$ of the counterfactual for each x_i in the *D=1* subsample which we base on observations of the *D*=0 subsample with values x_j being *close* to x_i. The conventional *one-to-one* matching estimator uses only the closest observation, which has given it the name "matching" estimator, as it matches observations from the one subsample to the other subsample.

In their seminal paper, Rosenbaum and Rubin (1983) have shown that independence conditional on *X*, see equation (3), implies independence conditional on *p(x)*, where $p(x) = \Pr(D=1 \mid X=x)$ is the one-dimensional *propensity score*. (In fact, Rosenbaum and Rubin (1983) assumed full independence instead of only mean independence as in (3). Frölich (2007) showed that mean independence (3) is nevertheless sufficient for consistency of propensity score matching.) Instead of finding observations with a similar value of *X*, it suffices to find observations with a similar value of *p(X)*. This reduces the dimensionality of the nonparametric estimation because *p(X)* is one-dimensional.

Alternatively, one could 'match' (i.e. find similar observations) on *p(X)* and a subset of *X*. Such combinations, which Rosenbaum and Rubin (1983) referred to as balancing scores, can help to ensure that a misspecification of the functional form of the propensity score has only a minor impact. We denote this subset as $\tilde{X}$, which usually contains variables that are suspected to be highly correlated with the outcome variable as well as with *D*.

The small sample properties of propensity score matching estimators have been explored and appear to be quite robust in different practical applications (e.g. Gerfin and Lechner, 2002, Gerfin et al. 2005, Larsson 2003). Moreover, it was subjected to several Monte Carlo studies investigating small sample problems and sensitivity issues (e.g. Frölich 2004, Lechner 2002). In this paper we use an extension of conventional matching estimation in several directions, similar to Lechner et al. (2006): First, as mentioned above, matching does not only proceed with respect to the propensity score but incorporates additionally some other covariates $\tilde{X}$. Second, instead of using first-nearest neighbour matching, all neighbours within a pre-specified radius are used. The estimator $\hat{m}_0(x)$ we use is essentially a kernel regression estimator

$$\hat{m}_o(x) = \frac{\sum_{j:D_j=0} y_j \cdot K_h\left(\left\|x_j - x\right\|\right)}{\sum_{j:D_j=0} K_h\left(\left\|x_j - x\right\|\right)}, \tag{5}$$

which uses only the *D*=0 observations. K_h is a kernel function and $\left(\left\|x_j - x\right\|\right)$ refers to a generalised distance between x_j and x. Since we match on the propensity score p and $\tilde{x}$ we define this distance as follows. Let $(p_j, \tilde{x}_j')'$ be the column vector of the estimated propensity score $\hat{p}(x_j)$ and $\tilde{x}_j$ of observation j. Then we can define the distance as:

$$\| x_j - x \| = (p_j - p(x), (\tilde{x}_j - \tilde{x})') \cdot W \cdot ((p_j - p(x), (\tilde{x}_j - \tilde{x})')\}, \qquad (6)$$

where *W* is a weight matrix. If *W*= $\boldsymbol{\Omega}^{-1}$, where $\boldsymbol{\Omega}$ is the empirical covariance matrix of $(p_j, \tilde{x}_j{}')$ in the *D*=0 subsample, we have the usual Mahalanobis distance. However, in our applications *W* is a modified version of $\boldsymbol{\Omega}^{-1}$ to ensure that the propensity score is not dominated by the covariates $\tilde{x}$. As suggested by Lechner et al. (2006), in W, weights implicitly assigned to the propensity score by $\boldsymbol{\Omega}^{-1}$ are multiplied by five.

We use a triangular kernel function $K_h(u) = h - u$ for $u < h$ and zero otherwise and where *h* is some specified radius. Thus, in (5) the kernel provides weights, which decrease with the distance of observations in the *D=0* group from *x*, and excludes observations that are more distant from *x* than *h*. The motivation for radius matching is the possibility of efficiency gains by comparing each treated individual with *several* similar non-treated individuals instead of just with *one* nearest neighbour without the risk of incurring too much additional bias. We aim to choose a rather small radius *h* to be more cautious with respect to bias than with respect to variance, because the variance of the estimator is visible after the estimation, whereas the bias generally is not. In particular, we use a conservative choice suggested by Lechner et al. (2006). If ∂ is the greatest distance between the set of *D=1* observations after eliminating those outside the common support and each ones corresponding closest neighbours in *D=0*, then this choice is $h = 0.9\partial$. This ensures that only those *D=0* observations are used, which have a distance to 'their' treated observation of no more than 90% of the worst match that we had obtained by one-to-one matching. If the neighbourhood defined by *h* contains zero observations for a particular *D*=1 observation, we use the first nearest neighbour in this case. The results are not very sensitive to the exact way the weighting is implemented. When a smaller value than 90% is used, the treatment effects change little but their estimated variances increase slightly.

As a third modification of the conventional one-to-one matching estimator, we also reduce bias by a weighted regression procedure recommended by Imbens (2004) and related to proposals by Rubin (1973) and Robins and Rotnitzky (1995). The working paper version of this paper provides further technical details (Behncke et al., 2007).

Another issue concerns the common support, which is investigated using the estimated propensity score based on a probit model: For observations of the *D=1* sample with propensity score *p(x)* very close to one, we may not be able to find a corresponding observation in the *D=0* sample with characteristics leading to similar values of *p(x)*. However, it will be seen in the following section that the common support is very large, and that the loss of observations is negligible.

The standard errors of ATET estimates modify the methods of Lechner et al. (2006) which are conditional on the weights for the comparison observations. This involves allowing for possible dependencies between unemployed persons due to clustering within caseworkers. Appendix B.1 and B.2 present more details for the estimators of the effect and the standard errors.

5 Analysis of the determinants of cooperativeness

5.1 Estimation of the propensity score

As a first step, we examine the determinants of cooperativeness of caseworkers allocated to the unemployed individuals in the sample. This is done with a probit regression where the dependent variable is the cooperativeness (or lack thereof) of the caseworker. Independent variables are characteristics of the unemployed person, of her caseworker, of the local labour market, and of the employment office. These estimates also serve as estimates of the propensity score for the subsequent impact analysis. Standard errors are adjusted to take into account clustering at the caseworker level.

To check the robustness of our results, we will consider four different sets of control variables *X*. In our main specification, we use *Xset 1* with 65 regressors. Table 4 shows its regression results. A first observation is that many of the coefficients are insignificant. This implies that caseworkers' attitudes and behaviour are more like a personal characteristic of the caseworker instead of merely mirroring an adaptation to the external environment. The finding that many variables are insignificant cautions against a model with too many *X* variables as they might simply be adding noise to the propensity score matching estimator. We further observe that caseworkers who face many unskilled unemployed or who work in offices that internally specialise by occupation tend to be less cooperative. The latter may be because specialization by occupation will lead to a better knowledge of the employment situation and vacancies in the particular industry. Another observation is that many of the interaction terms with the language region are significant. This may be particularly related to the language in which the written questionnaire was conducted since the translations from German to French and Italian may not have been able to pick up all the nuances of language. We therefore retain all these interaction terms as control variables as they are capturing important differences between the language regions of Switzerland. Goodness-of-fit statistics of the probit estimates of the propensity score, e.g. Efron's (1978) R^2 is 0.06, indicate some overall descriptive power with a substantial amount of randomness remaining.

A concern might be that *Xset 1* contains too few covariates to make the conditional independence assumption plausible. Furthermore, additional variables related to the outcome variable could increase precision. On the other hand, including too many variables also runs the risk of including endogenous control variables (i.e. those affected by the treatment variable) and/or reducing the common support region. They may also introduce more noise into the estimation of the propensity score.

Table 4: Probit estimates for prime age population (age 24-55, Xset 1)

Binary dependent variable: being a less cooperative caseworker

N = 100222		coefficient	std error	in *Xset* 4
Constant		-0.24	0.36	*
French speaking employment office	*	1.39	0.73	*
Italian speaking employment office	***	4.75	1.28	*
Characteristics of the unemployed person				
Female		-0.04	0.03	*
x French speaking region		-0.10	0.07	
x Italian speaking region		0.03	0.08	
Mother tongue other than German, French, Italian		-0.03	0.04	*
x French speaking region	*	0.10	0.06	
x Italian speaking region		0.05	0.07	
Qualification: unskilled	**	0.10	0.04	*
x French speaking region		-0.13	0.08	
x Italian speaking region		-0.04	0.09	
Qualification: semiskilled		0.04	0.05	*
x French speaking region		0.00	0.08	
x Italian speaking region		-0.07	0.17	
Qualification: skilled without accredited degree		0.02	0.05	*
x French speaking region	**	0.19	0.09	
x Italian speaking region	*	-0.28	0.17	
Qualification: skilled with degree (=reference category)				
Number of unemployment spells in last two years		0.01	0.01	
x French speaking region		-0.01	0.02	
x Italian speaking region	**	0.05	0.02	
Fraction of time employed in last years		0.00	0.03	
x French speaking region	**	-0.13	0.06	
x Italian speaking region		0.03	0.09	
Employability low		0.02	0.11	
x French speaking region		0.15	0.17	
x Italian speaking region		0.15	0.20	
Employability medium		0.00	0.10	
x French speaking region		0.02	0.14	
x Italian speaking region		0.04	0.19	
Employability high (=reference category)				
Local labour market characteristics				
Unemployment rate in canton		0.06	0.06	
x French speaking region		-0.18	0.12	
x Italian speaking region	**	-0.27	0.14	

Table 4 to be continued

Table 4 continued

		Coefficient	std error	in *Xset* 4
Allocation of unemployed to caseworkers (reference: at random)				
By industry		0.14	0.10	
x French speaking region		-0.06	0.20	
x Italian speaking region		-0.45	0.36	
By occupation	**	0.24	0.10	*
x French speaking region		0.16	0.21	
x Italian speaking region		-0.03	0.33	
By age		0.12	0.22	
By employability		-0.09	0.17	
By region		0.06	0.13	
Other		-0.05	0.15	
Characteristics of the caseworker				
Age		-0.01	0.01	*
x French speaking region	*	-0.02	0.01	*
x Italian speaking region	***	-0.06	0.02	*
Female		-0.04	0.10	*
x French speaking region		0.07	0.20	
x Italian speaking region		0.02	0.35	
Experience in employment office (tenure in years)		0.02	0.02	*
x French speaking region		-0.03	0.03	
x Italian speaking region		-0.07	0.05	
Own experience of unemployment		-0.04	0.10	
x French speaking region		-0.15	0.21	
x Italian speaking region		0.03	0.38	
Indicator for missing caseworker characteristics		-0.10	0.25	
Education: above vocational training	*	-0.20	0.11	*
x French speaking region		0.35	0.25	
x Italian speaking region		-0.39	0.41	
Education: tertiary track (university or polytechnic)		-0.20	0.14	
x French speaking region		0.32	0.27	*
x Italian speaking region		-0.35	0.48	
Education: vocational/apprenticeship (reference category)				
Special vocational training of caseworker		0.09	0.12	*
x French speaking region		0.29	0.36	
x Italian speaking region		0.42	0.35	

Note: Maximum Likelihood estimates of the probit model. The estimated coefficients of the linear predictor after transforming probabilities by the probit link function are given. ***, **, * indicate significant coefficients at the 1%, 5% and 10% level, respectively. Standard errors are clustered at the caseworker level. Most variables are interacted with French and Italian language region. (German is the reference language region.) The last column indicates variables included in *Xset* 4 with ∗.

In *Xset 2* (= 232 regressors) we added a large number of additional covariates. The additional variables of the unemployed person are age, civil status, children, and earnings in the last job. Furthermore, there are three dummies for education, three dummies for foreign language knowledge, and two dummies for the types of foreigners' work permit. To approximate the unemployed person's labour market history, it contains variables capturing the duration of unemployment in the last two years, the average wage in the last ten years, the total number of employment spells in the last ten years, the number of employment spells in the last five years, an indicator of having been out of labour force in the last five years, the fraction of time being employed and unemployed in the last ten years, and a dummy for having a zero contribution time to the unemployment insurance. Furthermore, it contains 16 occupation dummies, six industry dummies, and a dummy for looking for a part-time job. With regard to local labour market characteristics, additional variables are municipality size, and the cantonal unemployment rate. All these variables are interacted with indicators for the French and Italian language regions. The pension data also indicate the first month of contribution (since 1990), if ever contributed. We also include this variable together with interaction terms with being young/old, and foreigner/Swiss. These interaction terms roughly pick up in which year a foreigner migrated to Switzerland.

Most of these variables turned out to be insignificant in the estimation of the propensity score. Despite being insignificant, they still affect the calculation of the propensity score and can thus introduce noise into the matching estimator. By sequentially deleting insignificant variables in the probit model, we generated another *Xset 3* (= 94 regressors), in an attempt to reduce noise due to insignificant variables. In a general to specific approach, we eliminated covariates whose Wald-test did not suggest any explanatory power at the 5% level. We tested sequentially whether the coefficients for groups of variables (e.g. all education dummies or all qualification dummies) are zero, starting from those variables with the lowest significance levels. However, we retained all variables of *Xset 1* here, even if insignificant. Hence, *Xset 1* is a strict subset of *Xset 3*. Further eliminating sequentially all variables with insignificant F-test leads us to *Xset 4* with 46 regressors, which is the most parsimonious specification.

Table 5 shows various goodness-of-fit statistics of the probit regression for these different sets of *X* variables. The two parsimonious sets *Xset 1* (obtained by deliberate choice) and *Xset 4* (obtained by statistical variable choice) appear to be hardly worse than the two most complex specifications. Comparing *Xset 2* with *Xset 4*, adding almost 200 regressors increases Efron's R^2 by less than 2 percentage points and reduces the number of wrong predictions by less than 800 from more than 40,000. The additional variables thus mainly introduce noise.

Table 5: Goodness of fit measures for different Xsets for prime age population (age 24-55)

Regressors	Number of covariates	Log-Likelihood	Efron's R^2	NWP
Xset 1	65	-66303	0.058	41496
Xset 2	232	-65408	0.074	40097
Xset 3	94	-65807	0.067	41031
Xset 4	46	-66478	0.056	40832

Note: The number of observations for each *Xset* is 100222. Efron's R^2 (Efron, 1978) is a measure for residual variation. NWP is the number of wrong predictions. For a detailed discussion, see Amemiya (1981).

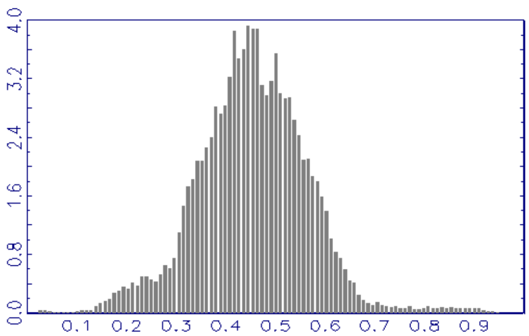
5.2 Common support

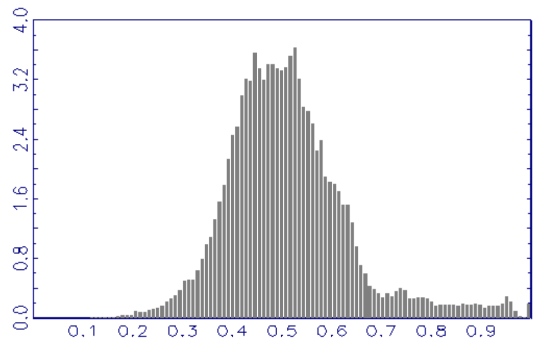
The nonparametric identification strategy relies on estimating the expected counterfactual outcome $E[Y^0 \mid X]$ for every *D=1* observation. This is possible only if for every value of *X* that is observed in the *D=1* population also at least one individual with very similar values of *X* can be found in the *D=0* population. Figure 3 shows histograms of the estimated propensity scores for the *D=1* and *D=0* subsamples.

Partly due to availability of a very large sample, the region of common support is large: Even very high values of the propensity score are observed in the *D=0* sample and also very small values are observed in the *D=1* sample. Observations that are outside of the common support are deleted: For estimating ATET, all *D=1* observations with a propensity score larger than the largest propensity score among all *D=0* observations are deleted. For estimating ATEN, all *D=0* observations with a propensity score smaller than the smallest propensity score among all *D=1* observations are deleted.

This leads to a loss of 312 treated observations (= 0.003%) for estimating ATET in our main specification with *Xset 1*. When estimating ATEN we lose 57 observations (= 0.0006%).

Figure 3: Distribution and common support of the propensity score





Note: Left graph is a histogram of the estimated propensity scores in the *D=0* sample. Right graph is a histogram of the estimated propensity scores in the *D=1* sample. Propensity score with *Xset 1*.

5.3 Matching quality

As suggested by Rosenbaum and Rubin (1983), it is desirable that after propensity score matching the joint and marginal distributions of *X* should be similar in the two matched subsamples. A basic test of this 'balance' is a test for the equality of the means. However, since we use a more complex estimation procedure for outcomes it seems more appropriate to apply a different procedure: we estimate the ATET as outlined for *Y* to each *X* as if it were the outcome. If balance were adequate, the expectation of these ATETs for each of the *Xs* would be zero and thus we test for these zero expectations.

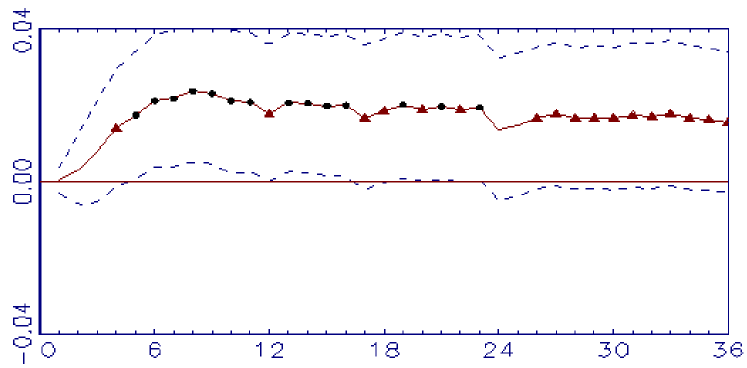
The procedure is applied to all variables from *Xset 1*. We find good matching quality in that almost all estimated treatment – control differences are quite small and statistically insignificant. The only exception is the number of unemployment spells in the past two years. In one of our later specifications, we therefore include this variable as an additional covariate on which matching is conducted.

6 Estimated treatment effects

6.1 Impact of a less cooperative caseworker

The following figures show the matching estimates when the propensity score is estimated with Xset 1. The distance function of the form (6) includes in $\tilde{x}$ the two dummy variables French speaking and Italian speaking employment office. They are included to ensure a high match quality with respect to these critical variables. Treatment is defined as *cooperation not very important* versus *very important*. That is, *D* is defined as one if the caseworker selected option 2 or option 3, and is defined zero if the caseworker selected option 1 (see Table 1). Thus, we compare caseworkers who place very much emphasis on cooperation to the rest.

Figure 4: Impact of having a less cooperative caseworker on employment in %-points



Note: Average treatment effect on the treated (ATET) on employment. Prime age unemployed (24 to 55 years). Abscissa: Month after registration of unemployment. Ordinate: Treatment effect on employment. Dots (triangles) indicate significance at the 5% (10%) level. The dashed line represents the point wise 95% confidence interval.

Figure 4 reports estimates for the outcome variable *employment* for the subsequent 36 months after registration. These results indicate that having a less cooperative caseworker increases the employment probability by about 2%-points. The effect sets in about five months after registration and is relatively stable until month 36. However, after the 24th month, it is only significant at the 10% level. Compared to the effects of other ALMP, which have often been found to be negative (Gerfin and Lechner, 2002; Gerfin et al., 2005; Lalive et al., 2006; Lalive et al., 2008), an effect of 2 percentage points is non-negligible.

A general concern with tougher caseworkers is that they might push jobseekers into precarious or unstable jobs, which due to the poor match quality might lead to higher job loss rates soon after. In order to examine the stability of jobs, we define an individual as being in *stable employment* in a given month if the employment spell is of at least a *six- (twelve-)* month duration without any interruption. Figure 5.1 gives the treatment effects on six-month stable employment period and Figure 5.2 gives them for twelve-month stable employment period. These figures show positive effects throughout that are slightly smaller and less significant than in Figure 4. Thus, we cannot rule out that non-cooperativeness of the caseworkers leads to less stable jobs.

Figure 5: Impact of having a less cooperative caseworker on stable employment

Figure 5.1: Six months stability

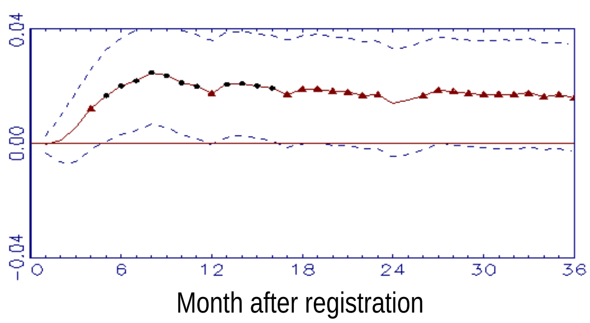
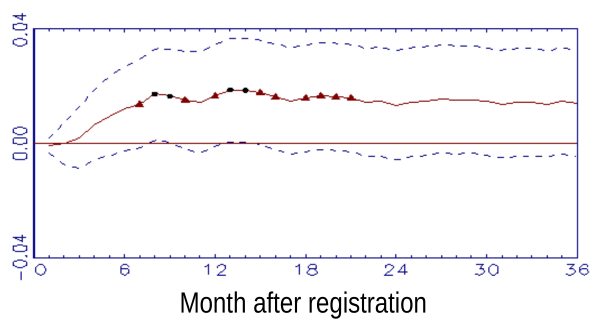


Figure 5.2: Twelve months stability



Note: Average treatment effect on the treated (ATET) on *stable* employment. In the left graph, employment is only considered as stable if the employment spell has a duration of at least 6 months. In the right graph, a duration of at least 12 months is required. Prime age unemployed (24 to 55 years). Abscissa: Month after registration of unemployment. Ordinate: Treatment effect on employment. Dots (triangles) indicate significance at the 5% (10%) level. Dashed line represents point wise 95% confidence interval.

6.2 Sensitivity analysis

We examine various alternative specifications to check the robustness of our results. The first three rows in Table 6 show effects of having a cooperative caseworker 6, 12, 18, and 30 months after registration of our main outcome specifications in terms of employment or stable employment as used in Figures 4 and 5. To examine effect heterogeneity, the next four rows show average treatment effects separately for four subgroups: qualified, unqualified, older than 55 years, and younger than 24 years. These estimates are insignificant, possibly due to the smaller sample in these age-subgroups, as opposed to the full sample. The subsequent rows examine alternative specifications of the estimator of the effect on employment outcomes and of the regressor set for the age group 24-55.

First, we split the sample according to the "criterion for allocation of unemployed to caseworkers", instead of including them as control variables *X* as done so far. When allocation is not at random, but based on observed characteristics of the unemployed such as his employability, we might be concerned that unobserved characteristics of the jobseeker may be correlated with the cooperation attitude of his caseworker, potentially leading to selection bias. However, when allocation is random, there is no such concern. We find that estimated treatment effects in both subgroups are quite similar to our main specification for employment in the first row. Furthermore, despite reduced sample size they are significant when allocation is random. This suggests that in our main specification selection on unobservables is unlikely to be present.

As another robustness check, we added cantonal dummies. There are 23 cantons in our estimation. 18 of the 22 cantonal dummies are insignificant in the probit estimation of the propensity score. The effects with cantonal dummies are slightly smaller than in Figure 4 and less precisely estimated.

Then, we reduce the radius from 0.9 to 0.1 in the propensity score matching estimator. This is similar to a reduction of the number of neighbours in k-nearest neighbour matching. Coefficients and standard errors are not affected much.

As an alternative specification, we exclude the regressor 'employability rating of the unemployed', which might potentially be endogenous as it is a subjective assessment made by the caseworker. For example, more demanding caseworkers might consider a certain type of client easier to place than the more lenient caseworkers would. The results do not change substantially. Second, we add the number of clients that caseworkers report to have counselled on average during the respective period to *Xset 1*, because one may want to control for caseworker's workload if more (or less) cooperative caseworkers are more successful in placing clients and hence more clients are assigned to them, which might negatively affect their efficiency. Again, the results are hardly affected. Third, we include a dummy for registering as unemployed in the second half of 2003 because in July 2003 the rules for benefit entitlement were tightened. Although it has no significant impact on cooperation behaviour, it reduces standard errors of employment effects by increasing estimation efficiency. Finally, we include the number of unemployment spells in the past two years as an additional covariate on which to match exactly, because the matching quality was imperfect with respect to this variable. The effects increase for month 6, but slightly decrease for the later months and become less significant.

The following rows present estimates of propensity score matching with *Xsets 2*, *3*, and *4*, respectively. They all have positive effects, but with different degrees of precision. It seems that the two very large regressor sets *2* and *3*, with many insignificant variables, lead to noisy estimates, whereas the results of regressor set *Xset 4*, where the insignificant variables have been purged, are more precise. Hence, including too many control variables that are not related to the treatment variable can introduce substantial noise into propensity score matching.

The ATEN estimates are positive but generally insignificant. (Only for month 30 is the estimate significant at the 10% level.) This suggests that those caseworkers who decided to be more demanding were right in doing so, whereas the gains from being more demanding are smaller or even zero for those caseworkers who decided against this strategy. This is what we would expect if caseworkers adapt to their environment.

The treatment effect estimates so far were obtained by propensity score matching as described in Section 4. As an alternative to this nonparametric estimator, we examined what the results would look like with a parametric approach. Thus, we estimated a Maximum Likelihood logistic regression of the outcome Y on a constant, D and X, separately for each month τ after registration. From these coefficient estimates we calculated the ATET by the change in the predicted probability that Y=1, if we change D from 0 to 1. This was averaged over all values of X that are observed in the D=1 subsample. As before, standard errors are clustered at the caseworker level. Furthermore, we use only those observations belonging to the common support as discussed in Section 5.2 to ensure that both approaches are based on the same population. The effects remain positive but smaller than those of the matching estimates. A potential explanation is that the logit estimate reflects some kind of average between ATET and ATEN. Whereas the nonparametric matching estimator does not restrict the form of the ATET in any way, such that it can be different to the ATEN if the distributions of X differ, the logistic functional form assumption strongly restricts it on X. Hence, the parametric estimates do not contradict the findings of a positive ATET, but their implicit assumptions on treatment effect heterogeneity appear to be too strong.

We also consider alternative definitions of the treatment variable. First, we discard caseworkers of the intermediate type and only consider an attitude as being less cooperative if the caseworker has chosen option 3. In the main specification, estimates are positive, but less precise due to the smaller sample size. Using the large *Xset 2*, estimates remain positive, but again less precise. The estimates of the parametric logit model are now similar to the matching estimates. Hence, in this scenario all estimates remain positive, but tend to decrease over time. The results for ATEN are even partly negative, although not statistically significantly so. We therefore restrict ourselves to interpreting the ATET estimates as overall positive, whereas not much can be concluded about the ATEN estimates.

In the last row of Table 6, we tested for possible non-linearities in the treatment variable. One could imagine that a caseworker with intermediate cooperation behaviour might perform better compared to a caseworker with very low or high cooperativeness. The respective estimates are close to zero and insignificant and therefore do not confirm this hypothesis.

Table 6: The impact of non-cooperativeness on employment, robustness analysis

	Obs.	Month 6 Effect	Month 6 t-Stat	Month 12 Effect	Month 12 t-Stat	Month 18 Effect	Month 18 t-Stat	Month 30 Effect	Month 30 t-Stat
		Effect at month after registration							
Non-cooperativeness (option 2 and 3) versus Cooperativeness (option 1)									
Alternative definitions of the outcome variables (Pscore matching with Xset 1, age 24-55)									
Employment	100222	0.021	2.360	0.018	1.955	0.019	1.917	0.017	1.725
Six-months stable employment	100222	0.020	2.316	0.018	1.908	0.019	1.959	0.017	1.785
Twelve-months stable employment	100222	0.012	1.614	0.017	1.825	0.016	1.656	0.015	1.532
Effects on employment for subgroups (Pscore matching with Xset)									
Qualified, age 24-55	61191	0.007	0.669	0.008	0.710	0.011	0.994	0.007	0.637
Unqualified, age 24-55	39031	0.017	1.384	0.007	0.552	0.003	0.229	0.000	-0.033
Older than 55 years	8580	-0.010	-0.626	0.012	0.690	-0.002	-0.115	0.009	0.437
Younger than 24 years	28980	-0.006	-0.408	0.005	0.310	0.016	1.091	0.009	0.615
Alternative specifications (age 24-55)									
PSM allocation at random (*Xset* 1)	55354	0.021	1.868	0.016	1.393	0.023	1.880	0.018	1.503
PSM allocation not at random (*Xset* 1)	43271	0.018	1.205	0.022	1.458	0.009	0.540	0.007	0.396
PSM with cantonal dummies (*Xset* 1)	100222	0.016	1.705	0.011	1.166	0.016	1.555	0.009	0.902
PSM with radius 0.1 (*Xset* 1)	100222	0.022	2.333	0.021	2.128	0.020	1.978	0.018	1.819
PSM without employability (*Xset* 1)	100222	0.018	2.047	0.020	2.247	0.016	1.658	0.016	1.614
PSM with number of clients (*Xset* 1)	100222	0.020	2.319	0.018	1.944	0.022	2.271	0.014	1.475
PSM with dummy for second half of 2003 (*Xset* 1) [a)]	100222	0.023	2.557	0.021	2.224	0.025	2.466	0.022	2.280
PSM exact on number of unemployment spells (*Xset* 1)	100222	0.023	2.611	0.016	1.784	0.018	1.922	0.014	1.514
Pscore matching (*Xset* 2)	100222	0.011	1.229	0.001	0.151	0.003	0.325	0.001	0.059
Pscore matching (*Xset* 3)	100222	0.010	1.137	0.001	0.083	0.004	0.368	-0.001	-0.056
Pscore matching (*Xset* 4)	100222	0.032	3.474	0.023	2.508	0.026	2.696	0.020	1.996
ATEN using PSM (*Xset* 1)	100222	0.002	0.028	0.002	0.209	0.008	0.813	0.016	1.680
Logit estimates (*Xset* 1)	100222	0.008	1.251	0.005	1.098	0.002	0.388	-0.002	-0.322
Non-cooperativeness (option 3) versus Cooperativeness (option 1), eliminating caseworkers with option 2									
Pscore matching (*Xset* 1, age 24-55)	60912	0.036	1.852	0.001	0.044	0.010	0.507	0.011	0.530
Pscore matching (*Xset* 2, age 24-55)	60912	0.032	1.534	0.004	0.232	0.009	0.438	0.011	0.530
ATEN using PSM (*Xset* 1, age 24-55)	60912	-0.038	-1.529	-0.012	-0.609	0.003	0.129	0.004	0.197
Logit estimates (*Xset* 1, age 24-55)	60912	0.028	2.583	0.008	1.182	0.012	1.290	0.010	1.010
Intermediate cooperativeness (option 2) versus no or full cooperativeness (option 1 or 3)									
Pscore matching (*Xset* 1, age 24-55)	100222	0.010	1.159	0.009	0.971	0.009	0.873	0.004	0.442

Note: Standard errors are clustered at the caseworker level. Pscore matching and PSM are abbreviations for propensity score matching.
[a)] PSM with dummy for registering as unemployed in the second half of 2003 (*Xset* 1)

6.3 Intermediate outcomes of caseworkers' cooperativeness

In this section, we examine the effects of the caseworker's attitude towards cooperation with his client on the use of instruments as potential intermediary channels of the total effect on employment. Two channels stand out: caseworkers may apply *sanctions* to different degrees and/or they make use of various active labour market *programmes* differently. We observe the number of realised sanction days and actual participation in programmes in the data. We do *not* observe whether caseworkers have threatened to use sanctions or to assign onerous programmes. The estimation process is analogous to Section 6.1 with the number of sanctions and programmes defined as the outcome variable. In particular, the sanction variable measures the total number of sanction days within the first year of registration, while programme participation is measured as a dummy variable taking the value 1 if the unemployed participated in at least one programme after registration of unemployment and until the end of 2006. We also examine a variant where we define programme participation as 1 if the unemployed person participated in at least two (or three, respectively) programmes, or in a particular type of programme.

From the first row in Table 7 we see that the effect on the total number of sanction days during the first 12 months is positive, but at 0.2 days quite small and not statistically significant. All the other results in Table 7 refer to the participation in ALMP. We find that the effects on the total number of ALMP and on the indicator whether the jobseeker participated in at least one, two or three programmes, respectively, are all extremely small and insignificant. The only thing we observe is that there may be a slight change in the composition of ALMP. When an ALMP is assigned, less cooperative caseworkers tend to choose *personality courses* more often, and *language skills training* or *computer skills training* less often. The effects nevertheless seem to be quite small, and are only weakly significant when we examine the first three ALMP together.

From a public unemployment insurance funds perspective, the consequences on the budget of these intermediate effects are small. A larger number of sanction days means that fewer benefits must be paid. The zero effect on the number of ALMP implies that no additional programme costs are incurred. The consequences of a very slight change in the composition of ALMP are somewhat unclear. Personality courses tend to be somewhat more expensive per day than language and computer training. However, personality courses are usually shorter than language training courses, but longer than computer training courses. Given the small change in composition, however, its financial consequences are probably small.

This weak evidence on the intermediate channels implies that the positive employment effects found in Figure 4 are mostly due to changes in behaviour or trust (including the threat of using sanctions or ALMP), and not due to an increased use of instruments such as sanctions or ALMP. Since the positive effects seem to last at for least three years, the total effect for the bottom line of unemployment insurance funds due to shorter unemployment spells, and less repeated unemployment is clearly positive.

Table 7: Intermediate outcomes of having a less cooperative caseworker

Effect on	Estimated treatment effect	t-Stat
Total number of sanction days in first 12 months after registration	0.186	0.940
Total number of ALMP participations after registration in 2003 (until end of 2006)	0.003	0.115
Participated in (at least **one**) ALMP after registration in 2003 (until end of 2006)	0.003	0.303
First programme after registration in 2003:		
Job search training	0.003	0.259
Personality courses	0.011	0.006
Language skills training	-0.007	0.005
Computer skills training	-0.006	-1.468
Vocational training	0.001	0.285
Employment programme or internship	0.000	0.099
Participated in at least **two** ALMP after registration in 2003 (until end of 2006)	-0.005	-0.657
Participated in at least **three** ALMP after registration in 2003 (until end of 2006)	0.000	-0.096
Within the **first three** programmes, participated at least once in		
Job search training	0.003	0.259
Personality courses	** 0.014	2.199
Language skills training	* -0.010	-1.784
Computer skills training	* -0.009	-1.864
Vocational training	0.000	0.042
Employment programme or internship	0.001	0.196

Note: ***, **, * indicates significance at 1%, 5% and 10% level, respectively. All programmes are observed after first registration in 2003 until the end of 2006.
Job search training is often short-term and provides participants with training in effective job search techniques. Personality courses help participants to position themselves in the labour market. Language skills training covers courses in foreign languages as well as alphabetization courses. Computer skills training includes mainly internet courses and office applications. Vocational training provides applicants with updated skills within their occupation. Employment programmes take place within a sheltered labour market. Internships and work in practice firms are also included in this category.

7 Conclusions

In most countries, caseworkers have substantial autonomy in the extent to which they cooperate with their clients. Some place more emphasis on counselling, whereas others consider monitoring of job search as their primary task. Using a linked jobseeker - caseworker data set, we investigate which attitude towards unemployed is more successful for their subsequent employment chances. These data allow us to control for potential selection bias by semiparametric matching estimators.

Our estimates suggest that the employment probabilities of those unemployed persons who were counselled by less cooperative caseworkers were higher due to their less cooperative attitude. They had about 2 percentage point higher employment probabilities during the first three years after registration than similar unemployed persons who were counselled by less demanding caseworkers. They also gained in terms of job stability. Positive employment effects did emerge mostly due to changes in behaviour and trust, and not due to increased use of sanctions or active labour market programmes. In an extensive sensitivity analysis, almost all results confirmed the sign of the effect, but in several cases, the effect was insignificant.

Regarding policy implementation, one may wonder how easy or difficult it would be to change a caseworker's cooperation attitude. External force may destroy caseworkers' intrinsic job motivation and could thus do more harm than good. However, caseworkers may be convinced that a very lenient attitude towards their clients may not help them. Moreover, when new caseworkers need to be hired, the cooperation attitude of the job applicants should be taken into consideration.

A Data Appendix

239,004 individuals registered as unemployed during the year 2003 at a Swiss employment office. We restrict our analysis to the 219,540 individuals registered at 103 regional employment offices that were responsible for a specific geographic area. We do not include the canton Geneva since its employment offices are functionally specialised. We exclude three offices that were newly established or re-organised during 2003, one employment office, which specialised on the difficult cases, and the tiny employment office in Appenzell-Innerrhoden, which did not participate in the survey. For 4,289 persons no caseworker was (yet) assigned, as it may take several weeks until an initial consultation with a caseworker takes place. We exclude foreigners without yearly or permanent work permits, because they are not entitled to all unemployment services. We also exclude individuals on disability or applying for it, and for the main analyses restrict the sample to the prime-age population. Finally, we lose about 25% of observations whose caseworker had either not responded to the questionnaire in general or the cooperativeness question in particular. Comparing the samples before and after dropping these observations, we do not find any significant differences in either their average characteristics or their average observed outcomes.

Table A.1: Sample selection criteria for empirical analysis

	Number of individuals	
	deleted	remaining
Population: all new jobseekers during the year 2003		239,004
Exclude Geneva and five other employment offices	-19,464	219,540
Exclude jobseekers not (yet) assigned to a caseworker	-4,289	215,251
Exclude foreigners without yearly or permanent work permit	-5,399	209,852
Exclude jobseekers without unemployment benefit claim	-18,434	191,418
Exclude jobseekers who applied for or claim disability insurance	-3,163	188,255
Restrict to prime-age population (24 to 55 years old)	-51,649	136,606
Exclude unemployed whose caseworker did not respond to the questionnaire	-31,469	105,137
Exclude unemployed whose caseworker did not respond to the cooperativeness question	-4,915	100,222

B Further details on the estimator

B.1 Implementation of the estimator

In this appendix, more details on the implementation of the estimation process are given. Table B.1 presents the matching algorithm.

Table B.1: A matching protocol for the estimation of ATET

Step 1	Estimate a probit model to obtain the choice probabilities: $\hat{P}_i = \Pr(D=1 \vert X = X_i)$
Step 2	Restrict sample to common support: Delete all *D=1* observations with $\hat{P}_i$ larger than the largest estimated propensity score among the *D=0* observations.
Step 3	*Estimate the counterfactual expectation of the outcome variable* $E[Y^0 \mid D=1]$
Step 3a	*Largest distance to first nearest neighbour* For every observation *i* in *D=1* find the nearest *D=0* observation in terms of distance, defined in (6) and denote it by ∂_i. Define the maximum distance of all ∂_i as ∂. This value ∂ would be a natural choice for the bandwidth h in the kernel function (5) as it ensures that for every *D=1* observation there is at least one *D=0* within a ∂ neighbourhood. Instead we use a smaller bandwidth $h = 0.9\partial$, because Lechner et al. (2006) recommend that this often leads to better finite sample properties, since the bias tends to dominate the variance. This implies that there are some *D=1* observations for which the neighbourhood defined by h does not contain a *D=0* observation. In this case, the bandwidth is extended locally such that the nearest *D=0* observation is included. (See also Heckman et al. 1998 on the need for asymptotic undersmoothing.)
Step 3b	*Radius matching step* With the choice of bandwidth *h*, calculate the ATET according to: $\frac{1}{N_1}\sum_{i:D_i=1}(Y_i - \hat{m}_0(X_i))$, where $\hat{m}_0(X_i)$ is a kernel function as given in (5).
Step 3c	*Exploit double robustness property* We follow Imbens (2004) to adjust for small mismatches in *X*, which occur when the matched comparison observation is not exactly identical in *X*, but only similar. More details are given in our discussion paper (Behncke et al., 2007).

Note: The table refers to the estimation of ATET. The modifications for ATEN are obvious. The vector $\tilde{x}$ includes the two dummy variables French speaking and Italian speaking employment office. $\tilde{x}$ is included to ensure a high match quality with respect to these critical variables, as has been explained in Section 4.2.

B.2 Standard errors for clustered matching

Using estimators of $m_0(X_i)$ of the form given by (5) it is readily seen that the ATET estimator may be expressed as the weighted average of Y_i in the form of $\sum_{i:D_i=1}\frac{Y_i}{N_i} - \sum_{i:D_i=0} Y_i w_i$,where $\sum_{i:D_i=0} w_i = 1$ and $\sum_{i:D_i=1}\frac{1}{N_i} = 1$. Following Lechner (2001), since the two terms are based on independent subsamples and treating the weights as given, the variance of ATET is the sum of the variances of these two terms. We consider first of all the second term and introduce the cluster structure by writing

$\sum_{i:D_i=0} Y_i w_i = \sum_{j=1}^{J} \sum_{i=1}^{N} I(C_i = j) \cdot (1-D_i) Y_i w_i$ where the sums are over all N unemployed in the sample and over the J caseworkers. The indicator function I takes on value unity, if C_i the caseworker assigned to person i is caseworker j. The number of clients of caseworker j in the D=0 sample and weighted by w_i is thus given by

$$N^j = \sum_{i=1}^{N} I(C_i = j)(1-D_i) w_i \ .$$

We now compute the variance by allowing that the outcomes across unemployed persons counselled by the same caseworker are dependent, but assume that observations across caseworkers are independent:

$$\begin{aligned} Var\left[\sum_{i:D_i=0} Y_i w_i\right] &= \sum_{j=1}^{J} Var\left[\sum_{i=1}^{N} I(C_i = j)(1-D_i) Y_i w_i\right] \\ &= \sum_{j=1}^{J} (N^j)^2 \, Var\left[\frac{1}{N^j} \sum_{i=1}^{N} I(C_i = j)(1-D_i) Y_i w_i\right] = \sum_{j=1}^{J} (N^j)^2 \, Var(A_j) \end{aligned}$$

where

$$A_j = \frac{1}{N^j} \sum_{i=1}^{N} I(C_i = j)(1-D_i) Y_i w_i \ .$$

Hence, the variance is obtained by summing over the caseworkers the variance of the expression A_j. Since the A_j are independent across the caseworkers we obtain

$$Var\left[\sum_{i:D_i=0} Y_i w_i\right] = Var(A) \sum_{j=1}^{J} (N^j)^2$$

and can estimate $Var(A)$ as

$$\widehat{Var}\left(A\right)=\frac{1}{J}\sum_{j=1}^{J}\left[A_j-\frac{1}{J}\sum_{j=1}^{J}A_j\right]^2,$$

which we now plug into the previous formula..

In the implementation, we ignore the regression step of the matching estimator, i.e. the estimation of the bias as described in Step 3c of Appendix B.1. The justification is given by Abadie and Imbens (2006), who demonstrated that a nonparametric regression step after the matching does remove the bias in the asymptotic distribution without affecting its variance. Although our estimator differs in some respects from the fixed-number-of-neighbours estimator they consider, the general set-up is very similar. It may be conjectured that our use of a parametric instead of a non-parametric regression may in fact *reduce* the variance, such that our inference would be conservative.

Acknowledgements

Markus Frölich is also affiliated with IZA, Bonn and IFAU, Uppsala. Michael Lechner has further affiliations with ZEW, Mannheim, CEPR, London, IZA, Bonn, and PSI, London. We are very grateful in particular to Heidi Steiger and Stephan Werner for helping us with the data and to Stephan Hammer for fruitful cooperation in this project. We also thank Martin Grubb, Thomas Ragni, an anonymous referee, the editor and the associate editor as well as seminar participants in Mannheim (ZEW), Nürnberg (Federal Employment Agency), St. Gallen (COST meeting), Oslo (EALE), and Munich (German Economic Association) for helpful comments and suggestions. All remaining errors are our own. We are grateful to the research fund of the Swiss unemployment insurance system (at the SECO) for providing the administrative database as well as substantial financial support for this project.

References

Abadie, A. and Imbens, G. W. (2006) Large sample properties of matching estimators for average treatment effects. *Econometrica*, **74**, 235-267.

Abbring, J., van den Berg, G. and van Ours, J. (2005) The effects of unemployment insurance sanctions on the transition rate from unemployment to employment. *Economic Journal*, **115**, 602-630.

Amemiya, T. (1981) Qualitative response models: A survey. *Journal of Economic Literature*, **19**, 1483-1536.

Ashenfelter, O., Ashmore, D. and Deschenes, O. (2000) Do unemployment insurance recipients actively seek work? Evidence from randomized trials in four U.S. States. *Journal of Econometrics*, **125**, 53-75.

Behncke, S., Frölich, M. and Lechner, M. (2007) Unemployed and their caseworkers - should they be friends or foes? *IZA working paper*, 3149.

Black D. A., Smith, J. A., Berger, M. and Noel, B. J. (2003) Is the threat of reemployment services more effective than the services themselves? Evidence from random assignment in the UI system," *American Economic Review*, **93**, 1313-1327.

Bloom, H., Hill, C. and Riccio, J. (2003) Linking program implementation and effectiveness: lessons from a pooled sample of welfare-to-work experiments. *Journal of Policy Analysis and Management*, **22**, 551-575.

Dolton, P. and O'Neill, D. (1996) Unemployment duration and the restart effect: some experimental evidence. *Economic Journal*, **106**, 387-400.

Efron, B. (1978) Regression and ANOVA with zero-one data: Measures of residual variation. *Journal of the American Statistical Association*, **73**, 113-212.

Fisher, R. A. (1935) *Design of Experiments*, Oliver and Boyd: Edinburgh.

Frölich, M. (2004) Finite sample properties of propensity-score matching and weighting estimators, *Review of Economics and Statistics*, 86(1), 77-90.

Frölich, M. (2007) Propensity score matching without conditional independence assumption - with an application to the gender wage gap in the UK, *Econometrics Journal*, 10, 359-407.

Frölich, M. (2008) Parametric and nonparametric regression in the presence of endogenous control variables, *International Statistical Review*, 76 (2), 214-227.

Frölich, M., Lechner, M., Behncke, S., Hammer, S., Schmidt, N., Menegale, S., Lehmann, A. and Iten, R. (2007) *Einfluss der Rav auf die Wiedereingliederung von Stellensuchenden*, Schweizerisches Staatssekretariat für Wirtschaft (SECO), SECO Publikation, Arbeitsmarktpolitik No 20, http://www.seco.admin.ch/dokumentation/publikation/00008/02015.

Gerfin M. and Lechner, M. (2002) Microeconometric evaluation of the active labour market policy in Switzerland. *Economic Journal*, **112**, 854-893.

Gerfin, M., Lechner, M. and Steiger, H. (2005) Does subsidised temporary employment get the unemployed back to work? An econometric analysis of two different schemes. *Labour Economics*, **12**, 807-835.

Gorter, C. and Kalb, G. R. J. (1996) Estimating the effect of counselling and monitoring the unemployed using a job search model. *Journal of Human Resources*, **31**, 590-610.

Graversen, B. and van Ours, J. (2006) How to help unemployed find jobs quickly; experimental evidence from a mandatory activation program, *Discussion Paper, Center*, Tilburg University.

Heckman, J., Ichimura, H. and Todd, P. (1998) Matching as an econometric evaluation estimator. *Review of Economic Studies*, **65**, 261-294.

Heckman, J., LaLonde, R. and Smith, J. (1999) The economics and econometrics of active labor market programs," in O. Ashenfelter and D. Card (eds.), *Handbook of Labour Economics*, **3**, 1865-2097.

Imbens, G. W. (2004) Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, **86**, 4-29.

Larsson, L. (2003) Evaluation of Swedish youth labor market programs. *Journal of Human Resources,* **38**, 891-927.

Lalive, R., van Ours, J. and Zweimüller, J. (2005) The effect of benefit sanctions on the duration of unemployment. *Journal of European Economic Association*, **3**, 1386-1407.

Lalive, R., Zehnder, T. and Zweimüller, J. (2006) Makroökonomische Evaluation der Aktiven Arbeitsmarktpolitik der Schweiz. mimeo, University of Zürich.

Lalive, R. and Zehnder, T. (2007) The effects of public employment programs on equilibrium unemployment. mimeo, University of Zürich.

Lalive, R., van Ours, J. and Zweimüller, J. (2008) The impact of active labour market programmes on the duration of unemployment in Switzerland. *Economic Journal*, **118**, 235-257.

Lechner, M. (2001) Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. in: M. Lechner and F. Pfeiffer (eds.), *Econometric Evaluation of Active Labour Market Policies*, 43-58, Heidelberg: Physica.

Lechner M. (2002) Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods. *Journal of the Royal Statistical Society, Series A*, **165**, 59-82.

Lechner, M., Miquel, R. and Wunsch, C. (2006) Long-run effects of public sector sponsored training in West Germany. revised version of *Discussion Paper 2004-19, Department of Economics*, University of St. Gallen.

Martin, J. and Grubb, D. (2001) What works and for whom: A review of OECD countries' experiences with active labour market policies. *Swedish Economic Policy Review*, **8**, 9-56.

Meyer, B. D. (1995) Lessons from the U.S. unemployment insurance experiments. *Journal of Economic Literature*, **33**, 91-131

Neyman, J. (1923) On the application of probability theory to agricultural experiments. Essay on principles. *Statistical Science*, Reprint, **5**, 463-480.

Robins, J. M. and Rotnitzky, A. (1995) Semiparametric efficiency in multivariate regression models with missing data, *Journal of the American Statistical Association*, **90**, 122–129.

Rogers, W.H. (1993) Regression standard errors in clustered samples, *Stata Technical Bulletin*, **13**, 19-23.

Rosenbaum P. and Rubin, D. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70**, 41-55.

Rubin, D. (1973) The use of matched sampling and regression adjustments to remove bias in observational studies. *Biometrics*, **29**, 185–203.

Rubin, D. (1974) Estimating causal effects of treatments in randomized and nonrandomized Studies," *Journal of Educational Psychology*, **66**, 688-701.

Rubin, D. (1979) Using multivariate matched sampling and regression adjustment to control bias in observational studies. *Journal of the American Statistical Association*, **74**, 318-328.

Van den Berg, G. J., van der Klaauw, B. and van Ours, J. (2004) Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics*, **22**, 211-241.

Wunsch, C. (2005) Labour market policy in Germany: Institutions, instruments and reforms since unification," *Discussion Paper, Department of Economics*, University of St. Gallen.