Why do tougher caseworkers increase employment? The role of programme assignment as a causal mechanism

Martin Huber*, Michael Lechner**, and Giovanni Mellace+

*University of Fribourg
**University of St. Gallen
+University of Southern Denmark

This draft: December 2014

Date this version has been printed: ***04 December 2014***

Abstract: Previous research found that less accommodating caseworkers are more successful in placing unemployed workers into employment. This paper tries to shed more light on the causal mechanisms behind this result using semiparametric mediation analysis. Analysing very informative linked jobseeker-caseworker data for Switzerland, we find that the positive employment effects of less accommodating caseworkers are not driven by a particularly effective mix of labour market programmes they use, but rather by other dimensions of the counselling process, possibly including threat effects of sanctions, pressure to accept jobs, and other factors related to the caseworker's counselling style.

Keywords: Unemployment, counselling style, active labour market policy, direct effects, indirect effects, causal mechanisms, causal channels, matching estimation

JEL classification: J64, J68, C21, C31.

We have benefitted from comments by Judea Pearl as well as seminar participants in Melbourne (University of Melbourne), Sydney (University of Sydney), Dunedin (University of Otago), the EALE conference in Ljubljana, the workshop on Applied Microeconomics in Odense, and the COMPIE conference in Rome. Financial support from the Swiss National Science Foundation, grant number SNF 100018_137769 for the project "Identifying causal mechanisms in empirical economics" is gratefully acknowledged. The usual disclaimer applies.
Addresses for correspondence: Martin Huber (Martin.Huber@unifr.ch), Department of Economics, University of Fribourg, Bd. de Pérolles 90, CH-1700 Fribourg; Michael Lechner (Michael.Lechner@unisg.ch, www.michael-lechner.eu), Swiss Institute for Empirical Economic Research (SEW), University of St. Gallen, Varnbüelstrasse 14, CH-9000 St. Gallen; Giovanni Mellace (giome@sam.sdu.dk), Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M.

** Michael Lechner is also affiliated with CEPR and PSI, London, CESIfo, Munich, IAB, Nuremberg, and IZA, Bonn.

1 Introduction

In most countries, caseworkers in unemployment offices have dual roles of *counselling*, which requires some trust between caseworkers and 'their' unemployed, and *monitoring*, which requires for example policing job search efforts. The importance individual caseworkers assign to these tasks may differ. Consequently, some caseworkers pursue a more dominating and demanding strategy, while others aim at a more cooperative relationship. In a recent paper, Behncke, Frölich, and Lechner (2010; BFL from now on) found for Switzerland that less cooperative caseworkers increase the reemployment chances of unemployed by approximately 2%-points.

However, depending on this attitude, caseworkers might differ in their assignment of jobs, active labour market programmes, and imposition of sanctions, in addition to more personal channels such as counselling style. Of course, any of these dimensions may in principle affect the employment prospects of their clients. Therefore, in order to allow policy makers and caseworkers to learn from such studies, information is required about the driving factors that make less cooperative caseworkers more effective, for instance to consider these factors in the training of caseworkers. It does not appear satisfactory to advise caseworkers "to be tougher" without knowing how and why this strategy works. Understanding the 'why' may on the one hand be necessary to successfully promote a particular strategy among caseworkers and on the other hand allows them to adjust their behaviour more granularly by including only those elements of the "tough strategy" in their own approach that are responsible for the positive overall effects. However, since BFL only assessed the total effect of counselling style which subsumes all these factors, their study could not provide any answers to why exactly those gains appear.

Such a setting is typical for the current microeconometric literature on policy or treatment evaluation (see Imbens and Wooldridge, 2009), which focuses on estimating the total effects of some intervention of interest, such as the average treatment effect on the treated (ATET). As in the evaluation of caseworker style, not only total effects appear to be of policy relevance in many empirical problems, but also the causal channels or mechanisms through which they operate. One would then ideally like to disentangle the direct effect of the treatment on the outcome as well as the indirect ones that run through one or more intermediate variables, so-called mediators, in our case for instance assignment to active labour market programmes. Causal mechanisms are, however, not easily identified. Even if the treatment or policy intervention was randomly assigned, this would not imply randomness of the mediators (see Robins and Greenland, 1992), which may be regarded as intermediate outcomes. Therefore, the total effect cannot be disentangled by simply controlling for mediators, because in general, this would entail selection bias due to variables related to both the mediator and the outcome (see Rosenbaum, 1984). The assessment of causal mechanisms is thus more challenging in terms of identification and estimation than standard treatment evaluation problems.

This paper uses the mediation framework to decompose the positive ATET of being assigned to a less cooperative caseworker in an unemployment agency on the employment probability of unemployed individuals into an "indirect" effect through the assignment of labour market programmes and a "direct" effect, which comprises all remaining channels. To this end, we reconsider the linked jobseeker-caseworker data set for Switzerland analysed in BFL. We aim at opening the black box of the total effect of cooperativeness by considering the participation in labour market programmes, which is the most expensive tool that caseworkers may use, as an explicit mediator of caseworkers' counselling style.

To deal with both treatment and mediator endogeneity in the identification of the direct and indirect effects on the treated, we make use of a sequential conditional independence assumption.[1] The latter requires (i) that the treatment "caseworker cooperativeness" is

[1] This assumption is somewhat weaker than the sequential ignorability assumption of Imai, Keele, and Yamamoto (2010) (and others), who aim at identifying direct and indirect effects on the entire population rather than the treated.

conditionally independent of potential programme states and potential outcomes under non-treatment given the observed covariates and (ii) that the mediator "programme participation" is conditionally independent of the potential outcomes under non-treatment given the covariates and the treatment. We use semiparametric radius matching (see for instance Huber, Lechner, and Wunsch, 2013) on the treatment propensity score given the covariates and the mediator to disentangle the ATET into its direct and indirect components. Our findings suggest that the total effect is mainly driven by channels other than programme participation which increase employment by roughly 1.5 percentage points the initial months, but the effect levels off over time. In contrast, the indirect effects are neither economically, nor statistically significant. It therefore seems that the success of less accommodating caseworkers is not driven by a more effective mix of active labour market programmes (ALMP), but rather operates through other dimensions of the counselling style that possibly include the threat of sanctions or the pressure to go to job interviews.

The evaluation of direct and indirect effects, also known as mediation analysis (see Baron and Kenny, 1986, for an early paper), is widespread in statistics, epidemiology, political sciences, and psychology. Recently, Robins (2003), Petersen, Sinisi, and Laan (2006), VanderWeele (2009), Imai, Keele, and Yamamoto (2010), Albert and Nelson (2011), and Imai and Yamamoto (2013), among others, consider rather general mediation models under sequential conditional independence of the treatment and the mediator. In economics, however, comparably few studies aim at assessing causal mechanisms (some make conjectures about possible channels based on the effect of the treatment on particular mediators, however, without estimating a complete mediation model), even though the number has recently been increasing. For example, Flores and Flores-Lagunes (2009) evaluate the direct earnings effect of the US Job Corps programme when controlling for work experience as mediator, while Huber (2013) assesses direct and indirect health effects (via employment) of the same programme. Simonsen and Skipper (2006) estimate the direct wage effect of motherhood in

Denmark by controlling for several mediators through which motherhood may have an influence on wages. Heckman, Pinto, and Savelyev (2013) investigate the channels, namely cognitive skills and personality traits, through which the Perry Preschool programme (an early childhood intervention in the US) positively influenced participants' outcomes later in life such as employment, income, and crime. All of these papers use more or less credible conditional independence assumptions and vary in the restrictiveness of the functional form restrictions imposed.[2]

Our contribution to the literature is threefold: On the empirical side, this appears to be the first study that evaluates the causal mechanisms underlying the placement success of caseworkers based on a well-defined mediation framework, which plausibly controls for confounders of both the treatment (caseworker rigour) and the mediator (programme participation) by using a rich data base of linked survey and administrative information of caseworkers and their clients. On the methodological side, we adapt the sequential conditional independence assumption of Imai, Keele, and Yamamoto (2010) (among others), proposed for assessing direct and indirect effects on the entire population, to the treated population, which allows weakening some of the restrictions. The only other study focussing on the treated rather than the total population we are aware of is Vansteelandt and VanderWeele (2012), who impose a different set of identifying assumptions. Finally, in contrast to the vast majority of mediation analyses that rely on tight parametric specifications, we use a semi-parametric propensity score matching estimation approach that is flexible in terms of functional form assumptions and the heterogeneity of the effects. We demonstrate that matching is particularly attractive for the identification of the direct and indirect effects on the treated as considered in

---

[2] Only very few papers consider identification of causal mechanisms under selection on unobservables. Powdthavee, Lekfuangfu, and Wooden (2013) and Frölich and Huber (2014), use distinct instruments to tackle treatment and mediator endogeneity. Yamamoto (2013) assumes an instrument for the treatment and invokes a latent ignorability assumption similar to Frangakis and Rubin (1999) w.r.t. the mediator, which allows controlling for endogeneity of the latter despite the absence of a second instrument. Flores and Flores-Lagunes (2010, 2013) derive bounds on the direct and indirect effects, which are applied in Bampasidou et al. (2014) to disentangle the channel degree attainment in evaluating the effect of the Job Corps training programme.

this paper. Essentially, one only needs to run two matching estimations on two different propensity scores - the conditional probability of treatment given covariates and the conditional probability of treatment given covariates and mediators - to obtain the direct and indirect average treatment effects on the treated.

The remainder of this paper is organized as follows. In Section 2 we outline the relevant aspects of the Swiss active and passive labour market policy. In Section 3, we set up the econometric framework that allows us to define direct and indirect effects and to discuss the necessary identifying assumptions. In addition, estimation based on propensity score matching is outlined. Section 4 presents the institutional setting as well as the data. In Section 5, selected descriptive statistics are provided and the plausibility of the identifying assumptions in this setting is discussed. Section 6 presents the results and Section 7 concludes. Appendix A contains the proof of a theorem derived in Section 3. Further information about the data and estimation is provided in Appendices B and C that are available on the internet.

2 The Swiss unemployment insurance system

2.1 Passive and active labour market policy

Switzerland used to experience very low unemployment rates before the recession of the mid 1990s, when unemployment rose to 5%. This triggered a comprehensive revision of the Federal Unemployment Insurance Act in 1996 which entailed the consolidation of the regional employment offices and an expansion and professionalization of the services provided. An important element of the revision was a change from a system with passive unemployment benefits to an active one with benefits (in principle) being conditional on participation in ALMP. As more broadly discussed in Gerfin and Lechner (2002), the latter can be grouped into three broad categories: (a) training courses, ranging from basic courses to specific work-related training, (b) employment programmes, which should not compete with regular firms but yet be as similar as possible to regular employment, and (c) temporary employment schemes with wage subsidies aimed at encouraging job seekers to accept job offers that pay less than their unemployment benefit by compensating the difference. Past evaluations in Gerfin and Lechner (2002) and Gerfin, Lechner, and Steiger (2005) showed the latter programmes to be particularly effective, while the effects of the other programmes appeared to be at least doubtful.

2.2 The role of caseworkers

In the Swiss system, the federal agency in charge, the State Secretariat for Economic Affairs (seco), sets targets which all employment offices and caseworkers should pursue. Major goals are rapid reemployment, avoidance of long-term unemployment, avoidance of benefit exhaustion, and avoidance of repeated unemployment. Every employment office computes annual measures that allow assessing its performance in the light of the targets. In the period under investigation, the performance of the individual office had no financial consequences but only affected its reputation. It is worth noting that there are no strict guidelines by the federal or cantonal governments as to how to reach these targets. Secondly, regional employment offices are quite autonomous in the implementation of the unemployment insurance law. Furthermore, caseworkers generally enjoy considerable leeway when dealing with their clients. Many regional 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 tight bureaucratic restrictions (for more details see BFL).

Table 2.1: Survey question on cooperativeness of the caseworker

How important do you consider the cooperation with the jobseeker, regarding placements 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 and ALMP should sometimes be assigned or declined in spite of unemployed person's wishes.
☐ 3	Cooperation is less important; I should assign placements and ALMP independent of the wishes of the unemployed person

Note: 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.

Caseworkers have to fulfil two major roles, firstly to help unemployed clients to search and find appropriate employment, and secondly to monitor whether the clients search thoroughly enough and are willing to take up reasonable job offers. Some caseworkers put more emphasis on their role as counsellor and aim for a trustful relationship, whereas other caseworkers see their policing role as being more important and are, thus, more dominating and demanding towards their clients. On this issue there is data available from a question in a caseworker survey which is shown in Table 2.1, see BFL for more details.

3 Econometric framework

3.1 Potential outcomes and different causal effects

We are interested in disentangling the effect of a binary treatment (D), e. g. less accommodating caseworkers, on various labour market outcomes (Y) into a direct effect and an indirect effect operating through a possibly multidimensional mediator (M), e.g. participation in various labour market programmes. To define the parameters of interest, we use the potential outcome framework, which has been used in the context of direct and indirect effects for instance by Rubin (2004), Ten Have et al. (2007), and Albert (2008). We denote by $Y(d)$ and $M(d)$ the potential outcome and the potential mediator states under treatment $d \in \{1,0\}$.[3] Furthermore, the (total) average treatment effect on the treated is denoted by

[3] By defining the potential outcomes and mediators we implicitly impose the Stable Unit Treatment Value Assumption (SUTVA).

$\Delta_{D=1} = E\left[Y(1) - Y(0) \mid D = 1\right]$. To disentangle this total effects into direct and indirect (through *M*) causal channels, we first rewrite the potential outcome as a function of both the treatment and the mediator: *Y(d)=Y(d,M(d))*. This allows formulating the direct ($\theta_{D=1}(d)$) and indirect effects ($\delta_{D=1}(d)$) on the total and treated populations, respectively:

$$\theta_{D=1}(d) = E\left[Y(1, M(d)) - Y(0, M(d)) \mid D = 1\right],$$
$$\delta_{D=1}(d) = E\left[Y(d, M(1)) - Y(d, M(0)) \mid D = 1\right], d \in \{1, 0\}.$$

Robins and Greenland (1992) named these parameters the pure/total direct and indirect effects, however considering the effects on the total population rather than the treated. Pearl (2001) refers to the same estimands as natural direct and indirect effects, and Flores and Flores-Lagunes (2009) as net and mechanism average treatment effects, respectively. Note that the ATET is the sum of the direct and indirect effects defined upon opposite treatment states:

$$\begin{aligned}
\Delta_{D=1} &= E\left[Y(1, M(1)) - Y(0, M(0)) \mid D = 1\right] \\
&= E\left[Y(1, M(1)) - Y(0, M(1)) \mid D = 1\right] + E\left[Y(0, M(1)) - Y(0, M(0)) \mid D = 1\right] \\
&= \theta_{D=1}(1) + \delta_{D=1}(0) \\
&= E\left[Y(1, M(0)) - Y(0, M(0)) \mid D = 1\right] + E\left[Y(1, M(1)) - Y(1, M(0)) \mid D = 1\right] \\
&= \theta_{D=1}(0) + \delta_{D=1}(1).
\end{aligned} \tag{1}$$

This can be easily seen by adding and subtracting $E[Y(0, M(1)) \mid D = 1]$ after the second and $E[Y(1, M(0)) \mid D = 1]$ after the fourth equality. Furthermore, the notation $\theta_{D=1}(d)$ and $\delta_{D=1}(d)$ points to possibly heterogeneous effects w.r.t. the treatment state, i.e., interaction effects between the treatment and the mediator.

In contrast to the vast majority of the literature, Vansteelandt and VanderWeele (2012) consider the identification of (natural) direct and indirect effects on the treated (rather than the entire population). Concerning the direct effect, they argue that focussing on the potential mediator under treatment, $M(1)$, appears to be natural reference for treated subjects when the choice of reference levels appears a priori hard to justify, because *M(1)* corresponds to the actually observed choice of the treated. In this case, the direct effect is defined as:

$$\theta_{D=1}(1) = E\left[Y(1, M(1)) - Y(0, M(1)) \mid D = 1\right] = E\left[Y - Y(0, M(1)) \mid D = 1\right]. \tag{2}$$

The direct effect represents the fraction of the total effect of the treatment which is not attributed to the mediator. Thus, it includes all the remaining causal channels.

The indirect effect on the treated then naturally arises as the difference between the ATET and $\theta_{D=1}(1)$, which by decomposition (1) corresponds to $\delta_{D=1}(0)$:

$$\delta_{D=1}(0) = \Delta_{D=1} - \theta_{D=1}(1) = E\left[Y(0, M(1)) - Y(0, M(0)) \mid D = 1\right]. \tag{3}$$

The indirect effect is the fraction of the total effect of the treatment which can be attributed to the mediator.

In this paper, we aim at identifying and estimating (2) and (3). However, it is obvious that the direct and indirect effects cannot be identified without further assumptions, because by the observation rule either $Y(1, M(1))$ or $Y(0, M(0))$ is known for any unit: $Y = D\,Y(1, M(1)) + (1-D)\,Y(0, M(0))$. As a further complication, $Y(0, M(1))$ is never observed. Therefore, identification hinges on the existence of exogenous variation in the treatment and the mediator as discussed in the next section.

3.2 Identifying assumptions

To identify the effects of interest, we impose (sequential) conditional independence of the treatment and the mediator (Assumptions 1 and 2), which requires that we observe all factors that are jointly related (i) with *D* and the potential outcome and the mediator(s) under non-treatment and (ii) with *M* and the potential outcome under non-treatment.[4] Furthermore,

[4] As discussed in Pearl (2014), sequential conditional independence as imposed in Assumptions 1 and 2 may be relaxed when allowing for different sets of observed factors to control for various confounding problems in the associations between *D, M,* and *Y*. Here, we assume the same set of observables in either assumption, which appears reasonable in the light of our application, see Section 5.

the common support restriction (Assumption 3) implies that for each treated unit, there exist non-treated comparisons in the population that are similar in terms of observed covariates and mediator values.

Assumption 1: $\{Y(0,m), M(0)\} \amalg D \mid X = x$ for all *m* and *x* in the common support.

Assumption 1 states that the joint distribution of the potential outcomes (for any potential value *m*) and mediators under non-treatment are independent of the treatment conditional on *X*. This rules out unobserved confounders affecting the treatment on the one hand and the potential outcome and/or mediator under *D=0* on the other hand, after controlling for the covariates.[5] Note that the conditional independence of $\{Y(0,m), M(0)\}$ implies that $Y(0,M(0)) \amalg D \mid X$ holds, too, and that $Y(0,M(0)) = Y(0)$ is simply the potential outcome under non-treatment.

Assumption 2: $Y(0,m) \amalg M \mid X = x, D = d$ for all *m, d,* and *x* in the common support.

By Assumption 2 the observed mediator is independent of the potential outcome under non-treatment (again, under any potential value *m*) conditional on the covariates[6] and the treatment.[7]

Assumption 3: $\Pr(D = 1 \mid M, X) < 1$.

Assumption 3 states that there is no combination of *M, X* that entails treatment receipt with probability one, because otherwise, no suitable non-treated comparison observations in terms of *M, X* would exist. To see the intuition of this restriction, first note that it implies the weaker condition $\Pr(D = 1 \mid X) < 1$, which is the conventional common support assumption for the

[5] As the potential outcomes/mediators under treatment remain unrestricted (and thus may be dependent on treatment even conditional on *X*), Assumption 1 is somewhat weaker than equation (4) in Imai, Keele, and Yamamoto (2010) invoked for the identification of direct and indirect effects in the entire population.

[6] Notice that the set of covariates in Assumption 1 and 2 need not be the same as in our empirical application.

[7] This is weaker than equation (5) in Imai, Keele, and Yamamoto (2010) who impose this restriction also on $Y(1,m)$.

identification of the ATET. However, by an application of Bayes' theorem, it also follows from Assumption 3 that $\Pr(M=m|D=1,X)<1$ (or if M is continuous, that the conditional density of M given $D=1$, X is smaller than one). It is this additional implication which allows for the identification of direct and indirect effects, because it implies that for each treated observation, non-treated units with comparable M can be found conditional on X.[8]

Theorem 1 shows that under our assumptions, the counterfactual parameters required for the evaluation of direct and indirect effects are identified.

Theorem 1: Under Assumptions 1 to 3, the following equalities hold:

$$E[Y(0,M(1))|D=1]= E[E[Y|M=m,X=x,D=0]|D=1], \tag{4}$$

$$E[Y(0,M(0))|D=1]= E[E[Y|X=x,D=0]|D=1]. \tag{5}$$

Thus, $\theta_{D=1}(1)$ and $\delta_{D=1}(0)$ are identified.

The proof of Theorem 1 is relegated to Appendix A.[9] From an applied perspective, directly controlling for the potentially high dimensional vectors M and X when estimating (4) and (5) may be undesirable due to the curse of dimensionality problem that plagues nonparametric estimation. However, Rosenbaum and Rubin (1983) show that one may instead condition on the treatment propensity scores $p_{mx}(m,x)=\Pr(D=1|M=m,X=x)$ and $p_x(x)=\Pr(D=1|X=x)$, respectively, as they balance the distributions of (M, X) and X, respectively. This has the practical advantage that the vector of conditioning variables is reduced to a single dimension and, thus, circumvents the curse of dimensionality, at least in the case when parametric models provide good approximations of these probabilities:

[8] Note that our identifying assumptions differ from those of Vansteelandt and VanderWeele (2012) who in contrast to Assumption 1 allow for confounding of the mediator (see their equation (1)), at the price of imposing somewhat stronger conditional independence assumptions on the potential outcomes (under both treatment and non-treatment). Interestingly, they, however, obtain similar identification results.

[9] The result in equation (5) is the usual identification result in for the ATET (see Imbens, 2004).

$$\theta_{D=1}(1) = E[Y \mid D=1] - E[E[Y \mid p_{mx}(M,X), D=0] \mid D=1]\ ,$$
$$\delta_{D=1}(0) = E[E[Y \mid p_{mx}(M,X), D=0] \mid D=1] - E[E[Y \mid p_{x}(X), D=0] \mid D=1].$$

The direct and indirect effects can therefore be conveniently evaluated by propensity score-based matching of non-treated observations to the treated sample (i) using the estimated $p_{mx}(m,x)$ to obtain an estimate of $\theta_{D=1}(1)$ and (ii) using the estimated $p_x(x)$ to get the ATET from which the direct effect is then subtracted to obtain an estimate of $\delta_{D=1}(0)$. The next section discusses the features of the matching estimator used here.

3.3 Estimators

We use radius matching on the propensity score with bias adjustment to estimate the direct effect on the treated, the ATET, and - by subtracting the former from the latter - the indirect effect on the treated.[10] Estimation is semi parametric in the sense that only the propensity scores $p_x(x)$ and $p_{mx}(m,x)$ are parametrically specified, while the models for the conditional expectations of the mediator(s) and the outcomes are unrestricted. Propensity score matching is therefore more robust than fully parametric methods in terms of model specification and flexibly allows for effect heterogeneity in *X* and *M*.

Specifically, we apply an estimator that takes into account the methodological considerations of Lechner, Miquel, and Wunsch (2011). Compared to standard nearest-neighbour matching this procedure is more precise because it incorporates the idea of radius matching (e.g. Dehejia and Wahba, 2002). Furthermore, the algorithm uses the initial matching weights in a second step of (weighted) regression adjustment, which has two advantages. Firstly, the estimator satisfies a so-called double robustness property, which implies consistency if either the matching step is based on a correctly specified selection model or the regression model is

[10] There are several estimators in the literature which can be adapted to estimate the direct and indirect effects on the treated. For example one can use the inverse probability weight base estimator in Huber (2013) or the nonparametric estimator in Imai et al. (2010).

correctly specified (e.g., Rubin, 1979; Joffe et al., 2004). Secondly, the regression adjustment should reduce small sample as well as asymptotic biases of matching. Huber, Lechner, and Wunsch (2013) investigate the finite sample properties of this radius matching with bias adjustment algorithm along with many other matching type estimators and find it to be highly competitive.

Concerning inference, Abadie and Imbens (2008) show that for standard matching (i.e. based on a fixed number of comparison observations) bootstrap-based inference may be invalid. However, the matching algorithm used in our analysis is smoother than the one studied by Abadie and Imbens (2008) because it is based on a variable number of comparisons and uses the regression adjustment. For this reason, the bootstrap is most likely a valid inference procedure in our context. However, the standard bootstrap, which randomly draws unemployed with replacement, may underestimate the standard errors, because it does not take into account the correlation between unemployed counselled by same caseworker. Therefore, we rely on a block bootstrap that resamples caseworkers (rather than unemployed) along with all 'their' unemployed therein to account for clustering at the caseworker level. To be more precise, inference is based on (i) bootstrapping caseworkers 999 times; (ii) computing the bootstrap t-statistics of the respective average effects in each of the samples (normalized by the estimated effect); and (iii) estimating the p-value as the share of absolute bootstrap t-statistics that are larger than the absolute t-statistic in the original sample (see for instance MacKinnon, 2006, for a discussion on bootstrapping symmetric statistics). This statistic is smoothed as suggested by Racine and MacKinnon (2007). Since the theoretical findings by Abadie and Imbens (2006) and the simulation results in Huber, Lechner, and Wunsch (2013) suggest that the estimator is asymptotically normally distributed, bootstrapping the potentially pivotal t-statistic (computed under the assumption that the weights obtained to compute the control group are non-stochastic; see Lechner, 2002) has the advantage of potentially providing so-called asymptotic refinements and thus improving inference. In addition, we also checked the bootstrap distribution of the estimated effects directly (quantile method). The results are similar (available on request).

4 Empirical implementation

This section describes the data and the selection of our estimation sample.

4.1 The study sample

The population underlying our study sample consists of all individuals who registered at Swiss regional employment offices anytime during the year 2003. Very detailed individual information is available from the databases of the unemployment insurance system and social security records, including (among others) nationality, qualification, education, language skills, experience, profession, position, and industry of last job, occupation and industry of desired job, and an employability rating by the caseworker. The data also contain information on registration and deregistration of unemployment, benefit payments and sanctions, participation in ALMP, and employment histories since January 1990 with monthly information on earnings and employment status. Regional (labour market relevant) characteristics such as the cantonal unemployment rate and city size were matched to the individual information. Finally, the administrative data were linked to a caseworker survey based on a written questionnaire that was sent to all caseworkers in Switzerland who were employed at an employment office in 2003 and were still active in December 2004 at the time the questionnaire was sent (for all details, see BFL). The questionnaire contained questions about aims, strategies, processes, and organisation of the employment office and the caseworkers, among them the treatment variable about caseworker rigour. The sample selection follows exactly the steps explained in BFL and is shown in Appendix B.1. In the final sample, there are 1,284 caseworkers and 100,120 unemployed individuals.

4.2 Definition of treatment, mediators, and outcomes

As already mentioned above (Table 2.1), the caseworker questionnaire contains a question on how important she considers cooperation with the client. When comparing the answers to this question to the responses to other items of the questionnaire we observe less cooperative caseworkers to state that they tended to assign active labour market programmes to apply pressure and to check their clients' availability for jobs.

As in the main specification of BFL, we define the treatment (*D*) to be one if the caseworker chose option 2 or 3 (i.e., is less accommodating) and zero if she answered with option one (i.e., is more accommodating). Although this cooperation attitude of a caseworker may vary between her clients, we expect a cooperative caseworker to be more cooperative to all her clients than a less cooperative caseworker. The mediators (*M*) are defined as the *first* participation in an ALMP within six months[11] after the beginning of the current unemployment spell. To this end, we categorize the ALMPs into six mutually exclusive groups, out of which five are different types of training courses – namely job search training, personality course, language skill training, computer training, and vocational training – and the final one is participation in an employment programme or internship. Together with non-participation in any ALMP, this definition entails seven possible mediator states. In the empirical implementation, we simply include six dummies for whether each of the ALMPs was the first programme an unemployed participated in 2003 (if any).

The average duration (in days) of each training programme is reported in Table 4.1.

Table 4.1: Average duration in days of first ALPM started in the first 6 months.

Type of active labour market programme	*Mean duration in days*
Job search	22.35

[11] By definition a mediator should occur before the outcome. For this reason, we restrict the mediation period to six months, have month seven with neither measurements of mediators nor outcomes, and leave the remaining 28 months for the outcomes. Six month can be roughly considered as the median of the time until the first ALMP, as 48.81% of jobseekers started their first ALMP within six months (conditional on having started one at all).

Personality	40.00
Language	68.22
Computer	24.37
Vocational	46.58
Employment	129.56
All programmes	35.99

Notice that only 4.71% of the jobseekers (4,712 individuals) started a second ALMP and only 0.37% of all the jobseekers (373) started the second ALMP before the end of the first one, for the third ALMP the numbers are even smaller 1.56% (1,566 individuals) and 0.05% (46 individuals), respectively. For this reason we do not consider multiple participations in our analysis.

Table 4.2 gives the participation frequencies of ALMP across treatment states. At least unconditionally, job seekers of cooperative caseworkers tend to participate less in language, informatics, and vocational training than job seekers of non-cooperative caseworkers.

Table 4.2: Participation frequencies of ALMP by treatment states

Type of active labour market programme	Cooperative caseworker	Less cooperative caseworker	Difference
Job search	12.81	13.26	-0.45**
Personality	2.11	2.32	-0.21**
Language	2.50	2.31	0.19**
Computer	1.73	1.31	0.42***
Vocational	1.71	1.58	0.13
Employment	0.93	1.03	-0.09
All programmes	21.79	21.81	-0.02

Note: The entries in this table are the shares of job seekers participating in a specific programme in the first six months of their unemployment spell. *, **, *** means statistically different from zero at the 10, 5, and 1% level, respectively. Standard errors are clustered at the caseworker level.

Our main outcome of interest (Y) is employment, but we also consider registered unemployment with benefit receipt and a dummy for looking for a job in our analysis. These binary outcomes are available on monthly bases until the end of 2006. An individual is considered as employed in a particular month if she has de-registered at the employment office and the exit state is known to be employment. As the mediator causally precedes the outcome in our mediation framework, we only consider employment states assessed in month 8 after caseworker assignment (i.e., at least one month after programme start) to avoid problems of reversed causality. Our set-up implies that at least 29 outcome periods (on a monthly base) are available for any observation: At the latest, individuals are assigned to a caseworker in the end of 2003 so that the mediators are measured from the beginning until the first half of 2004 and the outcomes from the second half of 2004 until the end of the observation window in 2006. We therefore estimate the direct and indirect effects on employment over a period of 1.5 to 3 years after caseworker assignment.

5 The selection processes

Our identification strategy relies on the ability to observe all confounding variables of the relationship of the outcome with both the treatment and/or the mediator (conditional on treatment). When using the very detailed linked caseworker-client dataset, it is essential to understand which factors affect (1) the cooperation attitude and rigour of the caseworker and (2) her placement of unemployed into the programmes. For Switzerland, the first issue has been discussed in depth by BFL, while the second selection problem received considerable attention in the active labour market programme evaluation of Gerfin and Lechner (2002), which was based on a similar data set. For the sake of brevity, we only restate some of their arguments and refer the interested reader to those papers for more details.

The selection into the treatment (caseworker rigour) depends on three processes: which types of caseworkers are hired, how caseworkers are allocated to the unemployed, and how their attitudes develop after having been trained and gained experience on the job. In that caseworkers' attitudes may be related to their general skills of finding jobs for their clients, we use caseworker characteristics as controls such as their age, gender, education, work experience, and experience of own unemployment. We also control for the criteria by which the unemployed are assigned to caseworkers, which are known from the questionnaire. A further aspect is that caseworkers not only differ in personality, but also in how they react to the 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 colleague who, for example, is responsible mainly for youth. Therefore, we include in the analysis a range of covariates characterizing the unemployed persons' employment histories and the local labour market.

Following the arguments of Gerfin and Lechner (2002) as well as of other evaluation studies of active labour market policies in countries similar to Switzerland (e.g. see Lechner and Wunsch, 2013), the same factors identified as controls for selectivity related to caseworker type are also expected to influence selection into the programmes. One reason for this is that in the Swiss case, the caseworker is all-important in programme allocation. The unemployed has little power to affect the caseworker's decision. Although the set of variables required to control for selection is likely to be very similar, the relevance of the various factors may of course differ substantially for the treatment and the mediator. Finally, it is worth noting that programme assignment usually takes place early in the unemployment spell, so that time-varying (or dynamic) confounding of the mediators due to changes in the relevant factors during the unemployment spell is supposedly only a minor issue, if any. The fact that we use a (short) six months window for entering a programme is also in favour of this argument.

Table 5.1 shows the means of selected confounding variables by treatment and programme participation status. For brevity, programmes are aggregated into one group, see Table B.2 in Internet Appendix B for an extended set of descriptive statistics. The numbers suggest that there is limited selection with respect to caseworker rigour, perhaps with the exception of regional aspects. Selection into the programmes appears to be much stronger and driven by a larger number of factors. The selectivity pattern largely matches our expectations. For example, having low skills, working in agriculture, and having a 'problematic' labour market history is associated with a higher probability of programme participation. In addition, it seems that having a male caseworker, a caseworker with longer tenure, and not living in the German speaking part of Switzerland is also positively associated with the mediators.

Table 5.1: Descriptive statistics of main variables by treatment and mediator status

		Caseworker attitude			Training participation	
	Coop. Mean	Less c. Mean	Diff.	Participant Mean	Nonpart. Mean	Diff.
Individual characteristics						
Female	0.45	0.43	0.02	0.44	0.46	-0.02***
Qualification: unskilled	0.21	0.24	-0.02**	0.23	0.22	0.01**
Qualification: skilled with degree	0.58	0.56	0.02**	0.57	0.59	-0.02***
Number of unemployment spells in last two years	0.56	0.60	-0.04*	0.61	0.45	0.16***
Fraction of time employed in last two years	0.80	0.79	0.00	0.79	0.81	-0.01***
Number of employment spells in last 5 years (/ 10)	0.12	0.12	0.00	0.13	0.10	0.03***
Previous job: Working in primary sector	0.09	0.09	0.00	0.10	0.06	0.03***
Working in tertiary sector	0.59	0.58	0.02	0.58	0.61	-0.03***
Unskilled worker	0.28	0.29	-0.01	0.29	0.27	0.02***
Characteristics of the caseworker						
Female	0.42	0.41	0.01	0.41	0.44	-0.04***
Age	45.20	43.50	1.69**	44.36	44.47	-0.11
Tenure in employment office in years	5.75	5.92	-0.17	5.86	5.73	0.13**
Local labour market characteristics						
Employment office in German speaking region	0.69	0.69	0.00	0.67	0.77	-0.09***
French speaking region	0.26	0.21	0.05*	0.25	0.20	0.05***
Italian speaking region	0.05	0.10	-0.05***	0.08	0.04	0.04***
Cantonal unemployment rate (in %)	3.70	3.76	-0.05	3.74	3.68	0.06***

Note: *,**,*** means statistically different from zero at the 10%,5%,1%, respectively.

The conclusions from the descriptive statistics are largely confirmed in both estimated propensity scores, i.e. the probabilities of having a less cooperative caseworker conditional on (i) the covariates only and (ii) the (same) covariates and participation in the various programmes. Estimation is based on flexibly specified probit models, which closely follow the specification of BFL. The results for the two propensity score models are very similar and can be summarized as follows (detailed results are reported in Table B.3 of the Internet Appendix B.3): First, many of the coefficients are not statistically significant, pointing to limited selection into caseworker rigour. Second, caseworkers who face many unskilled jobseekers or who work in offices that internally specialize by occupation tend to be less cooperative, i.e. more demanding. Finally, many of the coefficients of the interaction terms with the language region are significant. Part of the reason may be that the translation of the written questionnaire from German to French and Italian was not perfect.

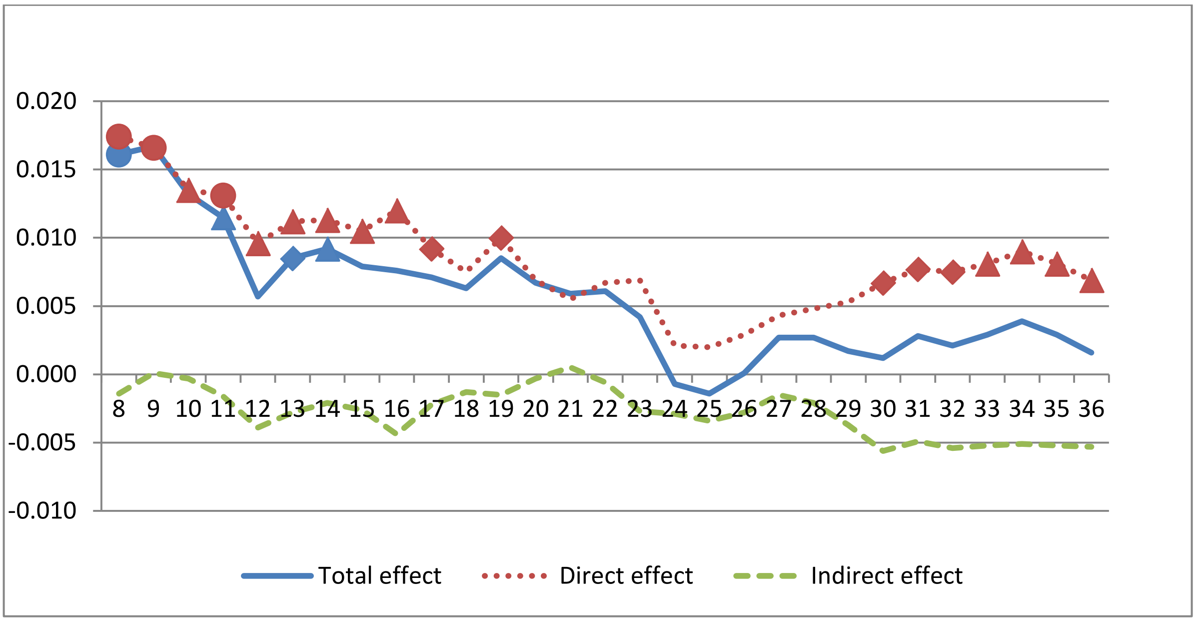
Internet Appendix C contains an extensive list of tests to check whether these propensity scores balance the characteristics of treated and non-treated in matching estimation (Tables C.1 and C.2). Balancing seems to work well, which suggests that the propensity scores are not subject to severe misspecification. The lack of strong selection is likely responsible for the fact that no serious common support issues occur, as not even 0.5% of the observations are off support (see Internet Appendix C for details).

6 Results

Figures 6.1 to 6.3 contain the ATET as well as the direct and indirect effects among the treated on three binary outcome variables, namely employment (Figure 6.1), registered unemployment with benefit receipt (Figure 6.2), and looking for a job (Figure 6.3), which (besides those receiving unemployment benefits) possibly includes unemployed without/with exhausted benefit claims. Estimates are provided on a monthly basis from month 8 to 36. The three lines in the tables represent the total, direct, and indirect effects and symbols superimposed on the respective lines indicate that these particular effects are (point wise) significant at the indicated level.

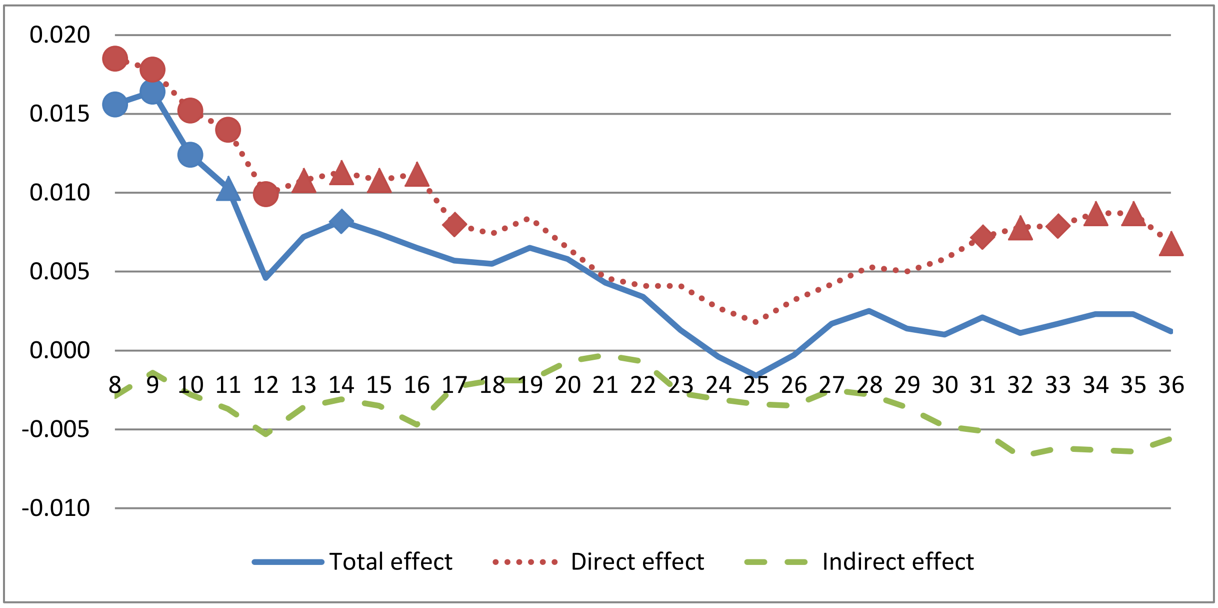
The results presented in Figure 6.1 suggest that less cooperative caseworkers significantly increase the reemployment probabilities of the unemployed in the beginning by roughly 1.5%-points. Over time, however, the ATET vanishes and is not statistically significantly different from zero anymore after month 14, even though it remains positive in almost all months. The (initial) employment gain is mainly driven by the direct effect of caseworker rigour, while the indirect mechanism through programme assignment seems to barely contribute to the ATET and is never significant. If anything, programme participation induced by less cooperative caseworkers might be detrimental in the long-run. This can be deduced from the fact that the direct effect is significant in the last months of the evaluation window, while the total effect is smaller and insignificant.

Figure 6.1: Effects on employment by month after registration



Note: A route / triangle / circle implies pointwise significance at the 10% / 5% / 1% level, respectively.

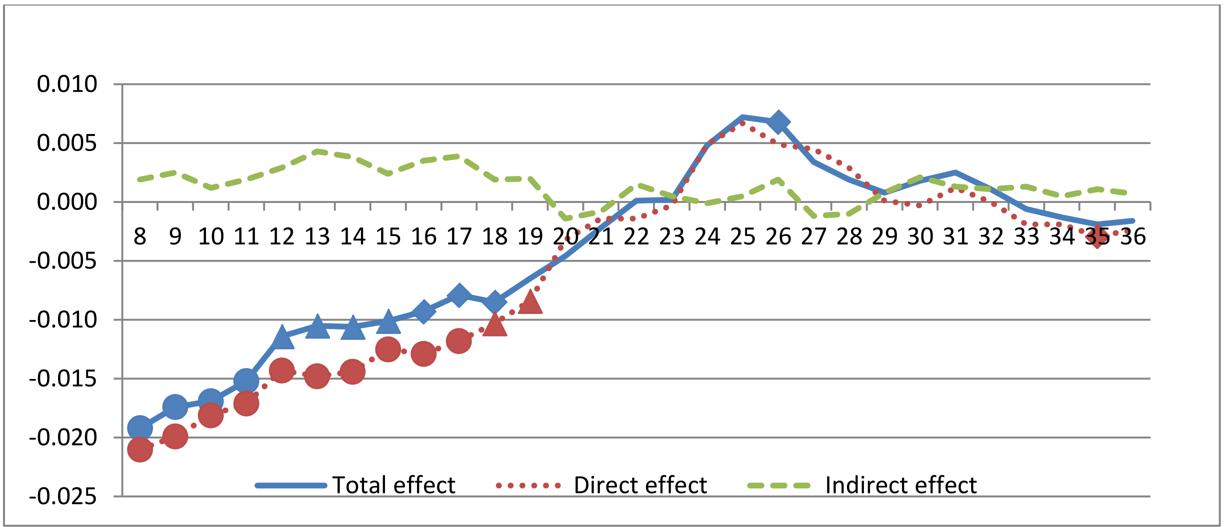
Figure 6.2: Effects on stable employment by month after registration



Note: A route / triangle / circle implies pointwise significance at the 10% / 5% / 1% level, respectively.

As pointed out by BFL, tougher caseworkers might push jobseekers into precarious or unstable jobs, which due to the poor match quality might lead to higher job loss rates soon after. However the effects on stable employment defined, as in BFL (the employment spell must be at least six month without interruption), are very similar and are reported in Figure 6.2.

Figure 6.3: Effects on unemployment by month after registration



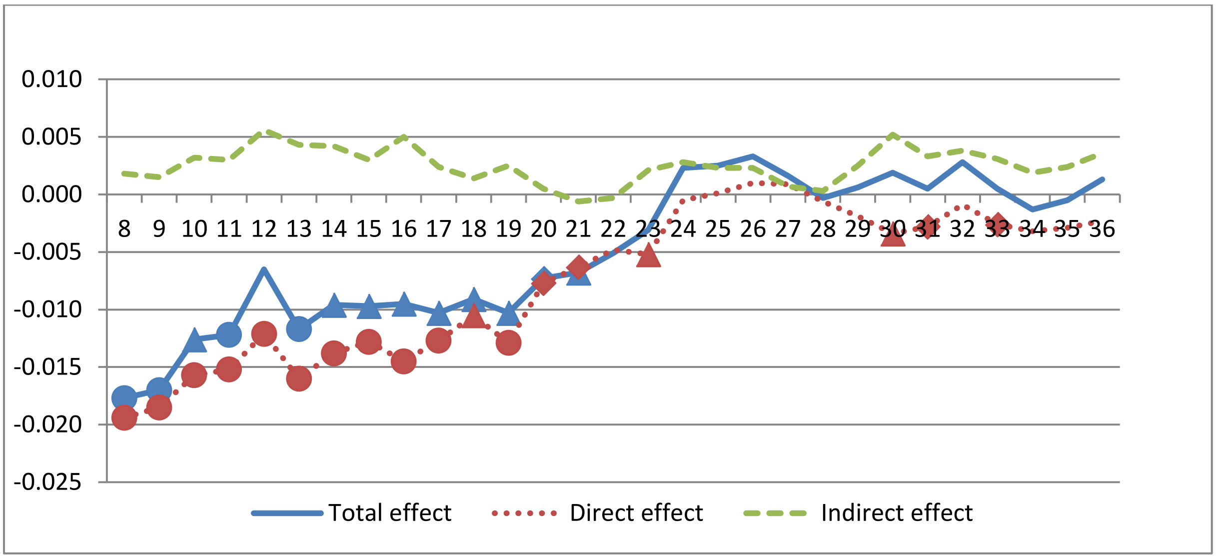
Note: A route / triangle / circle implies point wise significance at the 10% / 5% / 1% level, respectively.

Considering the unemployment and job search outcomes largely confirms the previous results (see Figures 6.3 and 6.4). Less cooperative caseworkers reduce registered unemployment and job search in initial periods by 1.5 to 2%-points, but the total effect levels off over time. Again, it is the direct effect which drives these findings, even though it is for either outcome now less pronounced at the end of the evaluation window than for employment. The indirect effect is again insignificant and close to zero in all months. Our estimates therefore suggest that the (at least initially) higher job placement and unemployment exit rates attained by less cooperative caseworkers are not driven by a better mix of active labour market programmes, but rather by other dimensions that possibly include the threat of sanctions[12] or the

[12] We do observe actual sanctioning days but as in BFL we do not find any significant effect of counselling style on sanctioning.

pressure to go to job interviews. The latter are however unobservable and thus cannot be investigated with in our data.

Figure 6.4: Effects on looking for a job by month after registration



Note: A route / triangle / circle implies point wise significance at the 10% / 5% / 1% level, respectively.

7 Conclusion

Considering Swiss labour market data, this paper for the first time decomposes the initially positive impact that rigorous caseworkers exert on their unemployed clients' employment perspectives into the (indirect) effect coming from the assignment to active labour market programmes and all other ("direct") causal channels, possibly including the threat of sanctions or the pressure to accept jobs. Using a sequential conditional independence assumption with respect to the treatment and the mediator, we estimate the direct and indirect employment effects on the unemployed clients of rigorous caseworkers by means of semiparametric radius matching on the propensity score. Our results suggest that the indirect effect is close to zero, implying that the success of rigorous caseworkers is not driven by assigning clients to particularly effective active labour market programmes. In contrast, the direct channels statistically significantly increase the employment probability by initially roughly 1.5%-points, but the effect levels off over time. Therefore, the success of rigorous caseworkers relative to more laid back colleagues in the initial months of our evaluation window seems to be caused by other factors of the counselling process rather than programme participation, possibly including the threat of sanctions or pressure to accept jobs. This suggests that policy makers should not only be interested in the effective provision of active labour market policies, but also in the analysis of other dimensions of caseworkers' counselling style, which can apparently make a difference.

References

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

Abadie, A., and Imbens, G.W. (2008): On the Failure of the Bootstrap for Matching Estimators. *Econometrica,* 76, 1537–1557.

Albert, J.M. (2008): Mediation analysis via potential outcomes models. *Statistics in Medicine*, 27, 1282–1304.

Albert, J.M., and Nelson, S. (2011): Generalized causal mediation analysis. *Biometrics*, 67, 1028–1038.

Bampasidou, M., Flores, C. A., Flores-Lagunes, A., and Parisian, D. (2014): The Role of Degree Attainment in the Differential Impact of Job Corps on Adolescents and Young Adults. *Research in Labor Economics,* forthcoming.

Baron, R.M., and Kenny, D.A. (1986): The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.

Behncke, S., Frölich, M., and Lechner, M. (2010): Unemployed and their caseworkers: should they be friends or foes? *Journal of the Royal Statistical Society Series A*, 173, 67–92.

Dehejia, R.H., and Wahba, S. (2002): Propensity-score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics,* 84, 151–161.

Flores, C.A., and Flores-Lagunes, A. (2009): Identification and Estimation of Causal Mechanisms and Net Effects of a Treatment under Unconfoundedness. IZA Dicussion Paper No. 4237.

Flores, C.A., and Flores-Lagunes, A. (2010): Nonparametric Partial Identification of Causal Net and Mechanism Average Treatment Effects. Mimeo.

Flores, C.A., and Flores-Lagunes, A. (2013): Partial Identification of Local Average Treatment Effects with an Invalid Instrument. *Journal of Business & Economic Statistics*, 31, 534–545.

Frangakis, C., and Rubin, D.B. (1999): Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika*, 86, 365–379.

Frölich, M. and Huber, M. (2014): Direct and Indirect Treatment Effects: Causal Chains and Mediation Analysis with Instrumental Variables. IZA Discussion Paper No. 8280.

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

Heckman, J., Pinto, R., and Savelyev, P. (2013): Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, 103, 2052–2086.

Huber, M. (2013): Identifying causal mechanisms (primarily) based on inverse probability weighting. *Journal of Applied Econometrics*, forthcoming.

Huber, M., Lechner, M., and Wunsch, C. (2013): The performance of estimators based on the propensity score. *Journal of Econometrics*, 175, 1–21.

Imai, K., and Yamamoto, T. (2013): Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments. *Political Analysis*, 21, 141–171.

Imai, K., Keele, L., and Yamamoto, T. (2010): Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, 25, 51–71.

Imbens, G.W. (2004): Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *The Review of Economics and Statistics*, 86, 4–29.

Imbens, G.W., and Wooldridge, J.M. (2009): Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47, 5–86.

Joffe, M.M., Ten Have, T.R., Feldman, H., and Kimmel, S.E. (2004): Model Selection, Confounder Control, and Marginal Structural Models: Review and New Applications. *The American Statistician*, *58*, 272–279.

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., and Wunsch, C. (2013): Sensitivity of matching-based program evaluations to the availability of control variables, *Labour Economics - An International Journal*, 21, 111–121.

Lechner, M., Miquel, R., and Wunsch, C. (2011): Long-Run Effects Of Public Sector Sponsored Training In West Germany. *Journal of the European Economic Association*, 9, 742–784.

MacKinnon, J.G. (2006): Bootstrap Methods in Econometrics. *The Economic Record*, 82, S2–S18.

Pearl, J. (2001): Direct and indirect effects. *In Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, San Francisco: Morgan Kaufman, 411–420.

Pearl, J. (2014): Interpretation and Identification of Causal Mediation. Forthcoming in *Psychological Methods*.

Petersen, M.L., Sinisi, S.E., and van der Laan, M.J. (2006): Estimation of Direct Causal Effects. *Epidemiology*, 17, 276–284.

Powdthavee, N., Lekfuangfu, W.N., and Wooden, M. (2013): The Marginal Income Effect of Education on Happiness: Estimating the Direct and Indirect Effects of Compulsory Schooling on Well-Being in Australia. IZA Discussion Paper No. 7365.

Racine, J.S., and MacKinnon, J.G. (2007): Inference via kernel smoothing of bootstrap P values, *Computational Statistics & Data Analysis*, 51, 5949–5957.

Robins, J.M. (2003): Semantics of causal DAG models and the identification of direct and indirect effects. In *Highly Structured Stochastic Systems*, P.J. Green, N.L. Hjort, and S. Richardson (eds.), Oxford: Oxford University Press, 70–81.

Robins, J.M., and Greenland, S. (1992): Identifiability and Exchangeability for Direct and Indirect Effects. *Epidemiology*, 3, 143–155.

Rosenbaum, P. (1984): The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of Royal Statistical Society, Series A*, 147, 656–666.

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

Rubin, D.B. (1979): Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. *Journal of the American Statistical Association*, 74, 318–328.

Rubin, D.B. (2004): Direct and Indirect Causal Effects via Potential Outcomes. *Scandinavian Journal of Statistics*, 31, 161–170.

Simonsen, M., and Skipper, L. (2006): The Costs of Motherhood: An Analysis Using Matching Estimators. *Journal of Applied Econometrics*, 21, 919–934.

Ten Have, T.R., Joffe, M.M., Lynch, K.G., Brown, G.K., Maisto, S.A., and Beck, A.T. (2007): Causal mediation analyses with rank preserving models. *Biometrics*, 63, 926–934.

VanderWeele, T.J. (2009): Marginal Structural Models for the Estimation of Direct and Indirect Effects. *Epidemiology*, 20, 18–26.

Vansteelandt, S.and VanderWeele, T.J. (2012). Natural direct and indirect effects on the exposed: effect decomposition under weaker assumptions. *Biometrics*, 68, 1019–1027.

Yamamoto, T. (2013): Identification and Estimation of Causal Mediation Effects with Treatment Noncompliance. Unpublished manuscript, MIT Department of Political Science.

Appendix A: Proof of Theorem 1

The proof considers the conditional expectations of the potential outcomes:

$$
\begin{aligned}
& E[Y(0,M(1)) \mid D=1] = \\
&\qquad = E[E[Y(0,m) \mid M(1)=m, X=x, D=1] \mid D=1] && \text{(law of iterated expectations)} \\
&\qquad = E[E[Y(0,m) \mid M=m, X=x, D=1] \mid D=1] && \text{(observation rule)} \\
&\qquad = E[E[Y(0,m) \mid X=x, D=1] \mid D=1] && \text{(Assumption 2)} \\
&\qquad = E[E[Y(0,m) \mid X=x, D=0] \mid D=1] && \text{(Assumption 1)} \\
&\qquad = E[E[Y(0,m) \mid M=m, X=x, D=0] \mid D=1] && \text{(Assumption 2)} \\
&\qquad = E[E[Y \mid M=m, X=x, D=0] \mid D=1]. && \text{(observation rule)} \\
& E[Y(0,M(0)) \mid D=1] = \\
&\qquad = E[E[Y(0,M(0)) \mid X=x, D=1] \mid D=1] && \text{(law of iterated expectations)} \\
&\qquad = E[E[Y(0,M(0)) \mid X=x, D=0] \mid D=1] && \text{(Assumption 1)} \\
&\qquad = E[E[Y \mid X=x, D=0] \mid D=1]. && \text{(observation rule)}
\end{aligned}
$$

It follows from the identification of $E[Y(0,M(1)) \mid D=1]$ and $E[Y(0,M(0)) \mid D=1]$ that $\theta_{D=1}(1)$, $\delta_{D=1}(0)$ are identified, too. To see this, note that by the observation rule, $E[Y(1,M(1))] = E[Y \mid D=1]$ so that the direct effect is given by:

$$\theta_{D=1}(1) = E[Y \mid D=1] - E[E[Y \mid X=x, D=0] \mid D=1]. \tag{4}$$

Furthermore, the indirect effect corresponds to

$$\delta_{D=1}(0) = E[E[Y \mid M=m, X=x, D=0] \mid D=1] - E[E[Y \mid X=x, D=0] \mid D=1]. \tag{5}$$

By decomposition (1) this is equal to $\Delta_{D=1} - \theta_{D=1}(1)$, where the ATET is given by $\Delta_{D=1} = E[Y \mid D=1] - E[E[Y \mid X=x, D=0] \mid D=1]$ due to Assumption 1. q.e.d.

Internet Appendix B: Data

Internet Appendix B.1: Selection of the sample

In this appendix, we describe how the estimation sample is selected. Since this is almost the same sample as used by Behncke, Frölich, and Lechner (2010),[13] we omit many details and refer the reader to this paper.

In total, 238,902 persons registered as new jobseekers during 2003. For each person we consider only the first registration in 2003 (while any further registrations may be regarded as outcomes of the caseworker contacts following the first registration). Each newly registered unemployed was linked to his first caseworker. Although, usually the same caseworker remains in charge for the entire spell of unemployment, we focus on the first caseworker to avoid any concerns about (rare but endogenous) caseworker changes. Furthermore, we exclude jobseekers who did not claim UI benefits as well as individuals who applied for or claim disability insurance, foreigners without permanent or yearly work permit, unemployed whose caseworkers are undefined, unemployed whose caseworkers did not answer the questionnaire (response rate: 84%) or did not answer the question on cooperativeness for some reason, and 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). The final sample size is 100,120 unemployed persons and is obtained by imposing the restrictions summarized in Table B.1.

[13] We detected and deleted 102 individuals who registered to the employment office before 2003 but were nevertheless included in the sample used by Behncke, Frölich and Lechner (2010).

Table B.1: Sample selection criteria for empirical analysis

Criteria		Individuals remaining
Population: all new jobseekers during the year 2003		238,902
Exclude Geneva and five other employment offices	-19'464	219,438
Exclude jobseekers not (yet) assigned to a caseworker	-4'289	215,149
Exclude foreigners without yearly or permanent work permit	-5'399	209,750
Exclude jobseekers without unemployment benefit claim	-18'434	191,316
Exclude jobseekers who applied for or claim disability insurance	-3'163	188,153
Restrict to prime-age population (24 to 55 years old)	-51'649	136,504
Exclude unemployed whose caseworker did not respond to the questionnaire	-31'469	105,035
Exclude unemployed whose caseworker did not respond to the cooperativeness question	-4'915	100,120

The initial number of caseworkers was 1,560. Only 15% of the initial caseworker did not answer the survey and only 42 caseworkers are missing after imposing the sample restrictions of Table B.1. This leaves us with 1,284 caseworkers.

Internet Appendix B.2: Descriptive statistics

Table B.2 shows the descriptive statistics by treatment and (aggregated) mediator group for the major variables used in the analysis.

Table B.2: Descriptive statistics of main variables by treatment and mediator status

		Caseworker attitude			Training participation		
		Coop. Mean	Less c. Mean	Diff.	Participant Mean	Nonpart. Mean	Diff.
Individual characteristics							
Female		0.45	0.43	0.02	0.44	0.46	-0.02***
Qualification: unskilled		0.21	0.24	-0.02**	0.23	0.22	0.01**
Qualification: semiskilled		0.16	0.16	0.00	0.16	0.15	0.01*
Qualification: skilled without degree		0.04	0.05	0.00	0.05	0.04	0.00
Qualification: skilled with degree		0.58	0.56	0.02**	0.57	0.59	-0.02***
Employability rating: low		0.13	0.14	-0.01	0.14	0.13	0.01*
Employability rating: medium		0.75	0.74	0.01	0.74	0.76	-0.03***
Employability rating: high		0.12	0.12	0.00	0.13	0.11	0.02***
Mother tongue other than German, French, Italian		0.31	0.32	-0.01	0.32	0.31	0.01***
Number of unemployment spells in last two years		0.56	0.60	-0.04*	0.61	0.45	0.16***
Fraction of time employed in last two years		0.80	0.79	0.00	0.79	0.81	-0.01***
Age		36.5	36.6	-0.10	36.4	37.0	-0.60***
Married		0.48	0.49	-0.01	0.48	0.49	-0.01*
Foreigner with B permit		0.13	0.14	-0.01**	0.14	0.14	0.00
Foreigner with C permit		0.24	0.25	-0.02***	0.25	0.22	0.03***
Lives in a big city		0.15	0.18	-0.03	0.16	0.16	0.00
Lives in a medium sized city		0.14	0.12	0.02	0.13	0.15	-0.02*
Past income		0.42	0.42	0.00	0.42	0.44	-0.03***
Number of employment spells in last 5 years (/ 10)		0.12	0.12	0.00	0.13	0.10	0.03***
Previous job:	Working in primary sector	0.09	0.09	0.00	0.10	0.06	0.03***
	Working in secondary sector	0.13	0.13	0.00	0.13	0.15	-0.02***
	Working in tertiary sector	0.59	0.58	0.02	0.58	0.61	-0.03***
Mother tongue in the canton's language		0.11	0.11	0.00	0.11	0.12	0.00
Previous job:	Self employed	0.01	0.01	0.00	0.01	0.00	0.00***
	Manager	0.08	0.07	0.00	0.07	0.08	-0.01***
	Skilled worker	0.61	0.6	0.00	0.61	0.61	-0.01
	Unskilled worker	0.28	0.29	-0.01	0.29	0.27	0.02***
Characteristics of the caseworker							
Female		0.42	0.41	0.01	0.41	0.44	-0.04***
Age		45.20	43.50	1.69**	44.36	44.47	-0.11
Tenure in employment office in years		5.75	5.92	-0.17	5.86	5.73	0.13**
Own experience of unemployment		0.65	0.62	0.03	0.63	0.64	-0.01
Education: vocational training		0.30	0.35	-0.06*	0.33	0.32	0.01
Education: above vocational training		0.46	0.41	0.05*	0.43	0.44	-0.01
Education: tertiary track (university or polytechnic)		0.24	0.24	0.01	0.24	0.24	0.00
Degree in vocational training for caseworkers		0.20	0.26	-0.06**	0.23	0.23	0.00
Indicator for missing caseworker characteristics		0.04	0.05	-0.01	0.04	0.04	0.00
Allocation of unemployed to caseworkers							
At random		0.22	0.23	-0.01	0.22	0.23	-0.01
By industry		0.52	0.57	-0.05	0.53	0.59	-0.05***
By occupation		0.52	0.61	-0.09	0.56	0.57	-0.02
By age		0.03	0.03	0.00	0.03	0.04	-0.01**
By employability		0.07	0.06	0.01	0.07	0.06	0.00
By region		0.12	0.12	0.00	0.12	0.10	0.02**
Other		0.08	0.07	0.01	0.07	0.08	0.00

Note: Table B.2 to be continued.

Table B.2 continued

	Caseworker attitude			Training participation		
	Coop. Mean	Less c. Mean	Diff.	Participant Mean	Nonpart. Mean	Diff.
Local labour market characteristics						
German speaking employment office	0.69	0.69	0.00	0.67	0.77	-0.09***
French speaking employment office	0.26	0.21	0.05*	0.25	0.20	0.05***
Italian speaking employment office	0.05	0.10	-0.05***	0.08	0.04	0.04***
Cantonal unemployment rate	3.70	3.76	-0.05	3.74	3.68	0.06***
GDP per capita in the canton	0.49	0.5	-0.01	0.49	0.50	-0.01***
Treatment	0	1	0.00	0.48	0.48	0.00
Mediators: training programmes						
Job search	12.81	13.26	-0.45**	0.60	0	
Personality	2.11	2.32	-0.21**	0.10	0	
Language	2.50	2.31	0.19**	0.11	0	
Computer	1.73	1.31	0.42***	0.07	0	
Vocational	1.71	1.58	0.13	0.08	0	
Employment	0.93	1.03	-0.09	0.60	0	
Number of caseworkers	670	614		1,240	1,282	
Number of unemployed	51,866	48,254		21,826	78,294	

Note: *, **, *** means statistically different from zero at the 10, 5, and 1% level, respectively. Standard errors are clustered at the caseworker level (1282 caseworkers). [a)] Multiple answers to this question were permitted. Hence, the means do not sum up to 1.

Internet Appendix B.3: Propensity score estimation

Table B.3 contains the result of a probit estimation of the treatment on our covariates with and without additional conditioning on the mediators.

Table B.3: Probit estimates of the propensity scores

Variable name	*P(x,m)* Marginal eff.	*P(x,m)* Std. error	*P(x)* Marginal eff.	*P(x)* Std. error
Characteristics of the caseworker				
Age	0.00	0.00	0.00	0.00
*French	-0.01*	0.00	-0.01*	0.00
*Italian	-0.02***	0.01	-0.02***	0.01
Female	-0.02	0.04	-0.02	0.04
*French	0.02	0.07	0.02	0.07
*Italian	0.01	0.13	0.01	0.13
Tenure in employment office (in years)	0.01	0.01	0.01	0.01
*French	-0.01	0.01	-0.01	0.01
*Italian	-0.02	0.02	-0.02	0.02
Own experience of unemployment	-0.01	0.04	-0.01	0.04
*French	-0.06	0.08	-0.06	0.08
*Italian	0.01	0.14	0.01	0.14
Indicator for missing caseworker characteristics	-0.04	0.09	-0.04	0.09
Education: above vocational training	-0.07*	0.04	-0.07*	0.04
*French	0.13	0.10	0.13	0.10
*Italian	-0.15	0.15	-0.15	0.15
Education: tertiary track (university or polytechnic)	-0.07	0.05	-0.07	0.05
*French	0.11	0.10	0.12	0.10
*Italian	-0.13	0.18	-0.13	0.18
Special vocational training of caseworker	0.03	0.04	0.03	0.04
*French	0.10	0.13	0.1	0.13
*Italian	0.16	0.13	0.16	0.13
Allocation of unemployed to caseworkers (ref. at random)				
By industry	0.06	0.04	0.06	0.04
*French	-0.02	0.08	-0.03	0.08
*Italian	-0.17	0.13	-0.17	0.13
By occupation	0.09**	0.04	0.09**	0.04
*French	0.06	0.08	0.06	0.08
*Italian	-0.01	0.13	-0.01	0.13
By age	0.05	0.09	0.05	0.09
By employability	-0.03	0.06	-0.03	0.06
By region	0.02	0.05	0.02	0.05
Other	-0.02	0.06	-0.02	0.06
Characteristics of the unemployed person				
Female	-0.01	0.01	-0.01	0.01
*French	-0.04	0.03	-0.04	0.03
*Italian	0.01	0.03	0.01	0.03
Mother tongue other than German, French, Italian	-0.03	0.02	-0.03*	0.02
*French	0.03	0.02	0.03	0.02
*Italian	0.02	0.03	0.02	0.03
Qualification: unskilled	0.03**	0.01	0.03**	0.01
*French	-0.04	0.03	-0.04	0.03
*Italian	-0.02	0.03	-0.02	0.03
Qualification: semiskilled	0.01	0.02	0.01	0.02
*French	0.00	0.03	0.00	0.03
*Italian	-0.03	0.06	-0.03	0.06
Qualification: skilled without degree	0.00	0.02	0.00	0.02
*French	0.07**	0.03	0.07**	0.03
*Italian	-0.10*	0.06	-0.10*	0.06

Table B.3 to be continued.

Table B.3 continued …

Variable name	P(x) Marginal eff.	P(x) Std. error	P(x,m) Marginal eff.	P(x,m) Std. error
Number of unemployment spells in last two years	0.00	0.00	0.00	0.00
*French	0.00	0.01	0.00	0.01
*Italian	0.02*	0.01	0.02*	0.01
Fraction of time employed in last two years	0.00	0.01	0.01	0.01
*French	-0.05**	0.02	-0.05**	0.02
*Italian	0.00	0.03	0.00	0.03
Employability low	0.01	0.04	0.01	0.04
*French	0.05	0.06	0.05	0.06
*Italian	0.05	0.07	0.05	0.07
Employability medium	0.00	0.04	0.00	0.04
*French	0.01	0.05	0.01	0.05
*Italian	0.02	0.07	0.02	0.07
$Age/10$	-0.01	0.02	-0.01	0.02
$Age^2/10000$	0.18	0.24	0.19	0.24
Married	0.01	0.00	0.01	0.00
Foreigner with B permit	0.03***	0.01	0.03***	0.01
Foreigner with C permit	0.02***	0.01	0.02***	0.01
Lives in a big city	0.03	0.04	0.04	0.04
Lives in a medium sized city	-0.03	0.03	-0.03	0.03
Past income	0.01	0.02	0.01	0.02
Number of employment spells in last 5 years (/ by 10)	0.03	0.03	0.03	0.03
Previous job: Working in primary sector	-0.02	0.02	-0.02	0.02
Working in secondary sector	0.00	0.02	0.00	0.02
Working in tertiary sector	-0.03*	0.02	-0.02	0.02
Mother tongue in the canton's language	0.01	0.01	0.01	0.01
Previous job: Self employed	0.01	0.03	0.01	0.03
Manager	0.01	0.02	0.01	0.02
Skilled worker	0.02	0.02	0.02	0.02
Unskilled worker	0.01	0.02	0.01	0.02
Local labour market characteristics				
Employment office in French speaking region	0.53**	0.27	0.53**	0.27
Employment office Italian speaking region	1.76***	0.49	1.76***	0.49
Unemployment rate canton	0.00	0.03	0.00	0.03
*French	-0.06	0.05	-0.06	0.05
*Italian	-0.08	0.06	-0.08	0.06
GDP per capita in the canton	0.18	0.29	0.17	0.29
Mediators: training programmes				
Job search	0.00	0.01	-	-
Personality	0.04*	0.03	-	-
Language	-0.02*	0.01	-	-
Computer	-0.05***	0.02	-	-
Vocational	0.00	0.02	-	-
Employment	0.04**	0.02	-	-

Note: Binary dependent variable: being a less cooperative caseworker. Model includes constant term. N = 100,200. Marginal effects are computed at the mean of each variable. Standard errors are clustered at the caseworker level (1282 caseworkers). *,**,*** means statistically different from zero at the 10%,5%,1% level, respectively.

Internet Appendix C: Common support and match quality

The balancing tests for the two propensity scores are reported in Tables C.1 and C.2, respectively. The first two columns contain the sample means for the treated and controls, respectively, while column 3 (bias) shows the difference of the two. In column (4) this difference is standardized by the average standard deviation (see Imbens and Woodridge, 2009, for discussions on this standardized bias, SB). The last two columns denote the t-statistic and its p-value of a two sample t-test for mean equality. The key take-away from these tables is that all standardized biases are comparatively small. The fact that some differences are significant nevertheless is driven by the very large sample size which will lead to even small differences becoming *statistically* significant.

Table C.1: Balancing tests (p-score: Treatment on covariates)

	Treated Mean	Control Mean	Bias	SB in %	t-value	p-value (%)
Characteristics of the Caseworker						
Age	43.62	43.28	0.35	2.26	3.50	0.05
*French	9.27	9.17	0.11	0.39	0.60	54.70
*Italian	3.78	3.20	0.58	4.15	6.43	0.00
Female	0.41	0.42	-0.01	-1.65	-2.56	1.05
*French	0.08	0.08	-0.01	-1.30	-2.02	4.37
*Italian	0.03	0.03	0.00	0.26	0.40	68.74
Tenure in employment office (in years)	5.92	5.96	-0.04	-0.97	-1.50	13.41
*French	1.33	1.35	-0.02	-0.45	-0.70	48.71
*Italian	0.65	0.59	0.07	2.49	3.85	0.01
Own experience of unemployment	0.61	0.61	0.01	1.00	1.55	12.02
*French	0.14	0.13	0.00	0.55	0.85	39.33
*Italian	0.06	0.04	0.02	6.51	10.09	0.00
Indicator for missing caseworker characteristics	0.41	0.42	-0.01	-2.06	-3.18	0.15
Education: above vocational training	0.09	0.10	-0.01	-1.63	-2.53	1.15
*French	0.03	0.03	0.00	-1.52	-2.35	1.88
*Italian	0.24	0.23	0.00	0.58	0.89	37.22
Education: tertiary track (university or polytechnic)	0.08	0.09	0.00	-0.18	-0.28	78.12
*French	0.02	0.02	0.01	3.63	5.63	0.00
*Italian	0.25	0.26	-0.01	-1.69	-2.62	0.89
Special vocational training of caseworker	0.02	0.02	0.00	-1.80	-2.79	0.53
*French	0.05	0.05	0.00	1.37	2.11	3.46
*Italian	0.04	0.05	-0.01	-1.90	-2.94	0.33
Allocation of unemployed to caseworkers (reference: at random)						
By industry	0.57	0.57	0.00	-0.22	-0.34	73.72
*French	0.08	0.09	0.00	-0.20	-0.30	76.29
*Italian	0.04	0.03	0.01	3.23	5.01	0.00
By occupation	0.61	0.60	0.01	0.80	1.24	21.62
*French	0.16	0.16	0.00	-0.03	-0.04	96.96
*Italian	0.06	0.05	0.00	1.30	2.01	4.45
By age	0.03	0.03	0.01	1.99	3.08	0.21
By employability	0.06	0.06	0.00	0.69	1.07	28.62
By region	0.12	0.11	0.01	1.29	2.00	4.57
Other	0.07	0.07	0.00	0.65	1.00	31.81
Characteristics of the unemployed person						
Female	0.43	0.43	0.00	0.00	0.00	100.00
*French	0.09	0.10	0.00	-1.05	-1.63	10.35
*Italian	0.04	0.04	0.01	2.85	4.41	0.00
Mother tongue other than German, French, Italian	0.32	0.32	0.00	0.03	0.04	96.76
*French	0.07	0.07	0.00	-0.63	-0.97	33.21
*Italian	0.03	0.02	0.00	2.20	3.40	0.07
Qualification: unskilled	0.24	0.23	0.00	0.13	0.20	84.23
*French	0.04	0.05	0.00	-0.63	-0.97	33.15
*Italian	0.03	0.03	0.01	2.76	4.27	0.00
Qualification: semiskilled	0.16	0.15	0.00	0.42	0.66	51.11
*French	0.04	0.04	0.00	-0.33	-0.50	61.43
*Italian	0.01	0.01	0.00	1.10	1.71	8.75
Qualification: skilled without degree	0.05	0.05	0.00	-0.43	-0.67	50.63
*French	0.02	0.02	0.00	-0.43	-0.67	50.35
*Italian	0.01	0.01	0.00	-0.60	-0.93	35.15

Note: Table C.1 to be continued.

Table C.1 continued …

		Treated Mean	Control Mean	Bias	SB in %	t-value	p-value (%)
Number of unemployment spells in last two years		0.59	0.61	-0.02	-1.00	-1.54	12.34
*French		0.15	0.15	0.00	-0.16	-0.25	80.58
*Italian		0.07	0.08	-0.01	-1.67	-2.58	0.99
Fraction of time employed in last two years		0.79	0.79	0.00	0.11	0.17	86.40
*French		0.16	0.16	0.00	-0.81	-1.26	20.94
*Italian		0.07	0.06	0.01	4.34	6.72	0.00
Employability low		0.14	0.15	-0.01	-1.34	-2.07	3.82
*French		0.02	0.02	0.00	-1.14	-1.76	7.86
*Italian		0.01	0.01	0.00	-2.27	-3.51	0.05
Employability medium		0.74	0.73	0.00	0.73	1.13	25.76
*French		0.17	0.17	0.00	-0.69	-1.06	28.80
*Italian		0.05	0.04	0.01	4.78	7.40	0.00
$Age/10$		3.66	3.65	0.01	0.62	0.96	33.77
$Age^2/10000$		0.14	0.14	0.00	0.62	0.96	33.87
Married		0.49	0.49	0.00	0.31	0.47	63.65
Foreigner with B permit		0.14	0.15	0.00	-0.29	-0.45	65.54
Foreigner with C permit		0.25	0.25	0.00	0.26	0.40	68.59
Lives in a big city		0.18	0.17	0.01	1.77	2.75	0.60
Lives in a medium sized city		0.13	0.13	-0.01	-1.88	-2.91	0.37
Past income		0.42	0.42	0.00	0.20	0.30	76.08
Number of employment spells in last 5 years		0.12	0.12	0.00	0.23	0.36	72.22
Previous job:	Working in primary sector	0.09	0.09	0.00	0.55	0.85	39.75
	Working in secondary sector	0.13	0.13	0.00	0.14	0.22	82.61
	Working in tertiary sector	0.58	0.58	0.00	-0.33	-0.51	61.12
Mother tongue in the canton's language		0.12	0.12	0.00	-0.49	-0.76	44.48
Previous job:	Self employed	0.01	0.01	0.00	0.02	0.03	97.92
	Manager	0.07	0.07	0.00	0.15	0.22	82.27
	Skilled worker	0.61	0.61	0.00	-0.54	-0.83	40.66
	Unskilled worker	0.29	0.29	0.00	0.52	0.80	42.32
Local labour market characteristics							
French speaking employment office		0.21	0.22	-0.01	-0.90	-1.39	16.48
Italian speaking employment office		0.09	0.08	0.01	4.34	6.72	0.00
Unemployment rate canton		3.75	3.73	0.03	2.38	3.68	0.02
*French		0.88	0.89	-0.01	-0.44	-0.67	50.05
*Italian		0.38	0.33	0.05	4.04	6.26	0.00
GDP per capita in the canton		0.50	0.50	0.00	-1.00	-1.55	12.09
Mediators: training programmes							
Job search		0.13	0.13	0.00	0.12	0.18	85.92
Personality		0.02	0.02	0.00	2.28	3.54	0.04
Language		0.02	0.03	0.00	-1.68	-2.60	0.93
Computer		0.01	0.02	0.00	-2.45	-3.80	0.02
Vocational		0.02	0.02	0.00	-0.21	-0.32	74.86
Employment		0.01	0.01	0.00	0.55	0.84	39.88

Note: N^1 = 47,978, N^0 = 51,779. SB denotes the standardized bias (see Imbens, 2004). Balancing tests are based on weights obtained by combining radius matching with weighted regression.

Table C.2: Balancing tests (p-score: Treatment on mediators and covariates)

	Treated Mean	Control Mean	Bias	SB in %	t-value	p-value (%)
Characteristics of the Caseworker						
Age	43.62	43.22	0.41	2.68	4.16	0.00
*French	9.27	9.07	0.20	0.75	1.16	24.54
*Italian	3.78	3.27	0.51	3.65	5.65	0.00
Female	0.41	0.42	-0.01	-1.92	-2.97	0.30
*French	0.08	0.08	0.00	-0.97	-1.50	13.34
*Italian	0.03	0.03	0.00	-0.40	-0.61	54.10
Tenure in employment office (in years)	5.92	5.96	-0.04	-0.99	-1.53	12.62
*French	1.33	1.33	0.00	-0.02	-0.03	97.98
*Italian	0.65	0.60	0.05	1.91	2.95	0.32
Own experience of unemployment	0.61	0.60	0.01	1.58	2.44	1.45
*French	0.14	0.13	0.01	1.63	2.53	1.14
*Italian	0.06	0.04	0.02	6.47	10.02	0.00
Indicator for missing caseworker characteristics	0.41	0.43	-0.02	-2.66	-4.12	0.00
Education: above vocational training	0.09	0.10	-0.01	-1.75	-2.71	0.68
*French	0.03	0.03	-0.01	-2.80	-4.33	0.00
*Italian	0.24	0.24	0.00	0.51	0.79	42.77
Education: tertiary track (university or polytechnic)	0.08	0.08	0.00	0.18	0.27	78.50
*French	0.02	0.02	0.01	3.29	5.10	0.00
*Italian	0.25	0.27	-0.01	-2.39	-3.71	0.02
Special vocational training of caseworker	0.02	0.02	0.00	-2.15	-3.33	0.09
*French	0.05	0.05	0.00	0.46	0.72	47.38
*Italian	0.04	0.05	0.00	-1.59	-2.47	1.37
Allocation of unemployed to caseworkers (reference: at random)						
By industry	0.57	0.57	0.00	0.26	0.40	69.24
*French	0.08	0.08	0.00	0.03	0.04	96.46
*Italian	0.04	0.03	0.01	2.55	3.94	0.01
By occupation	0.61	0.60	0.01	1.53	2.37	1.80
*French	0.16	0.16	0.00	0.63	0.98	32.74
*Italian	0.06	0.05	0.00	0.80	1.24	21.55
By age	0.03	0.03	0.00	1.33	2.05	4.00
By employability	0.06	0.06	0.00	0.27	0.41	67.89
By region	0.12	0.11	0.01	1.94	3.00	0.27
Other	0.07	0.07	0.00	0.00	0.00	99.73
Characteristics of the unemployed person						
Female	0.43	0.43	0.00	0.00	0.00	100.00
*French	0.09	0.09	0.00	-0.83	-1.29	19.87
*Italian	0.04	0.04	0.01	2.57	3.98	0.01
Mother tongue other than German, French, Italian	0.32	0.32	0.00	-0.43	-0.67	50.59
*French	0.07	0.07	0.00	-0.74	-1.15	25.03
*Italian	0.03	0.02	0.00	1.31	2.03	4.26
Qualification: unskilled	0.24	0.24	0.00	-0.15	-0.24	81.20
*French	0.04	0.05	0.00	-0.48	-0.74	45.66
*Italian	0.03	0.03	0.00	2.52	3.90	0.01
Qualification: semiskilled	0.16	0.16	0.00	0.19	0.30	76.80
*French	0.04	0.04	0.00	-0.19	-0.30	76.78
*Italian	0.01	0.01	0.00	0.73	1.12	26.11
Qualification: skilled without degree	0.05	0.05	0.00	-0.28	-0.43	67.06
*French	0.02	0.02	0.00	0.32	0.50	61.89
*Italian	0.01	0.01	0.00	-1.37	-2.12	3.42

Note: Table C.2 to be continued.

Table C.2 continued …

	Treated Mean	Control Mean	Bias	SB in %	t-value	p-value (%)
Number of unemployment spells in last two years	0.59	0.62	-0.02	-1.48	-2.29	2.21
*French	0.15	0.15	0.00	-0.21	-0.33	74.08
*Italian	0.07	0.08	-0.01	-1.85	-2.87	0.41
Fraction of time employed in last two years	0.79	0.80	0.00	-0.12	-0.19	85.11
*French	0.16	0.16	0.00	-0.44	-0.68	49.61
*Italian	0.07	0.06	0.01	3.72	5.77	0.00
Employability low	0.14	0.15	-0.01	-1.34	-2.07	3.81
*French	0.02	0.02	0.00	-1.14	-1.77	7.73
*Italian	0.01	0.01	0.00	-1.75	-2.71	0.68
Employability medium	0.74	0.73	0.01	1.13	1.75	8.08
*French	0.17	0.17	0.00	-0.08	-0.12	90.22
*Italian	0.05	0.04	0.01	4.86	7.53	0.00
$Age/10$	3.66	3.65	0.01	0.67	1.04	30.06
$Age^2/10000$	0.14	0.14	0.00	0.55	0.85	39.52
Married	0.49	0.49	0.00	-0.04	-0.06	95.53
Foreigner with B permit	0.14	0.15	0.00	-0.57	-0.88	38.12
Foreigner with C permit	0.25	0.25	0.00	0.14	0.22	82.36
Lives in a big city	0.18	0.17	0.01	1.46	2.26	2.38
Lives in a medium sized city	0.13	0.13	-0.01	-1.93	-2.99	0.28
Past income	0.42	0.42	0.00	0.38	0.60	55.16
Number of employment spells in last 5 years	0.12	0.13	0.00	-0.19	-0.29	77.26
Working in primary sector	0.09	0.09	0.00	0.28	0.43	66.84
Working in secondary sector	0.13	0.13	0.00	0.27	0.42	67.64
Working in tertiary sector	0.58	0.58	0.00	0.03	0.05	95.92
Mother tongue in the canton's language	0.12	0.12	0.00	-0.93	-1.44	14.90
Self employed	0.01	0.01	0.00	-0.17	-0.26	79.74
Manager	0.07	0.07	0.00	0.32	0.50	61.71
Skilled worker	0.61	0.61	-0.01	-0.90	-1.40	16.31
Unskilled worker	0.29	0.29	0.00	0.56	0.87	38.53
Mediators: training programmes						
Job search	0.13	0.13	0.00	0.00	0.00	100.00
Personality	0.02	0.02	0.00	0.00	0.00	100.00
Language	0.02	0.02	0.00	0.00	0.00	100.00
Computer	0.01	0.01	0.00	0.00	0.00	100.00
Vocational	0.02	0.02	0.00	0.00	0.00	100.00
Employment	0.01	0.01	0.00	0.00	0.00	100.00
Local labour market characteristics						
French speaking employment office	0.21	0.22	0.00	-0.36	-0.55	58.09
Italian speaking employment office	0.09	0.08	0.01	3.67	5.69	0.00
Unemployment rate canton	3.75	3.73	0.03	2.29	3.54	0.04
*French	0.88	0.88	0.00	0.09	0.14	88.70
*Italian	0.38	0.34	0.04	3.35	5.19	0.00
GDP per capita in the canton	0.50	0.50	0.00	-0.95	-1.47	14.06

Note: N^1 = 47,978, N^0 = 51,779. SB denotes the standardized bias (see Imbens, 2004). Balancing tests are based on weights obtained by combining radius matching with weighted regression.

Figures C.1 to C.4 provide histograms of the distribution of the respective propensity scores among the treated and controls in order to understand common support issues.

The graphs do not suggest any evident overlap problem among the outcome supports of treated and control. Overall, the number of observations on common support is 99’758 (99.64% of all observations).

Figure C.1: Propensity score distribution (without mediators) among the treated

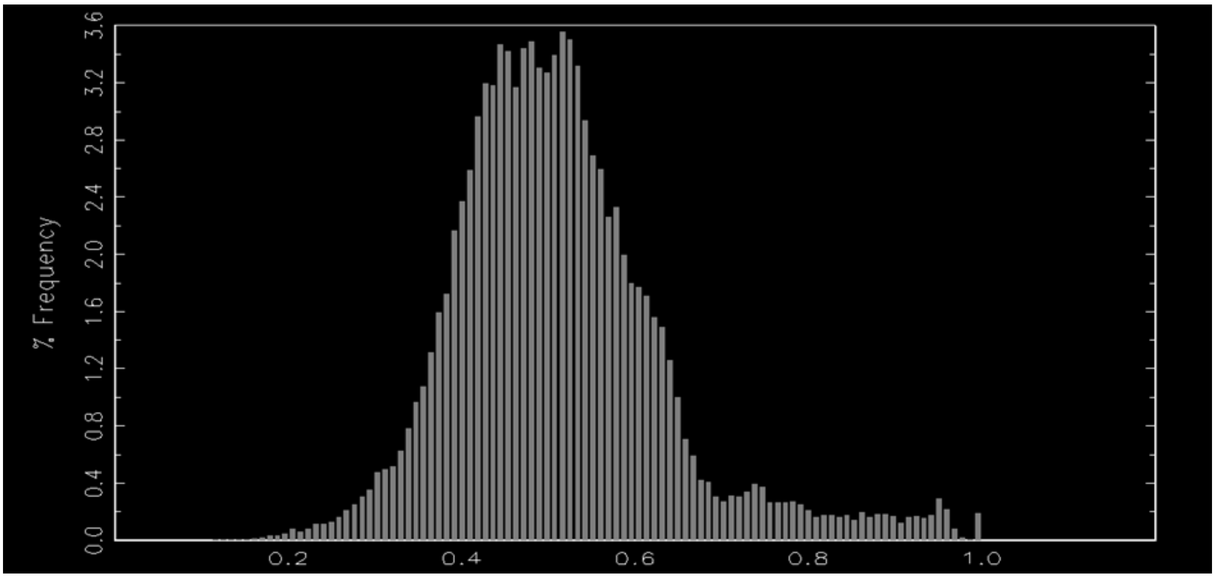


Figure C.2: Propensity score distribution (without mediators) among the controls

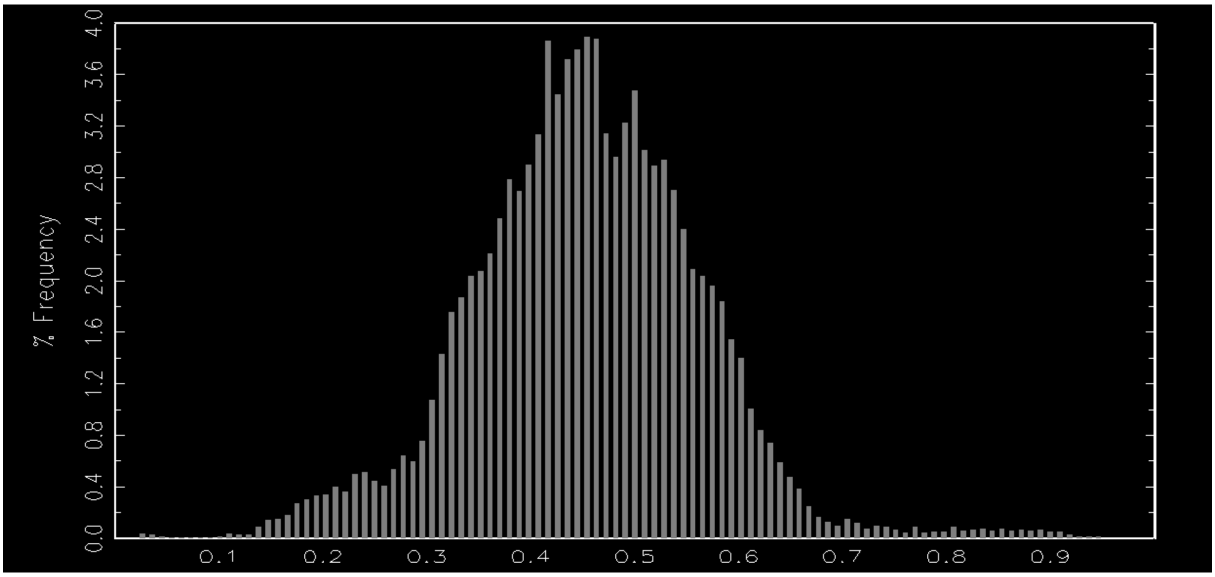


Figure C.3: Propensity score distribution (with mediators) among the treated

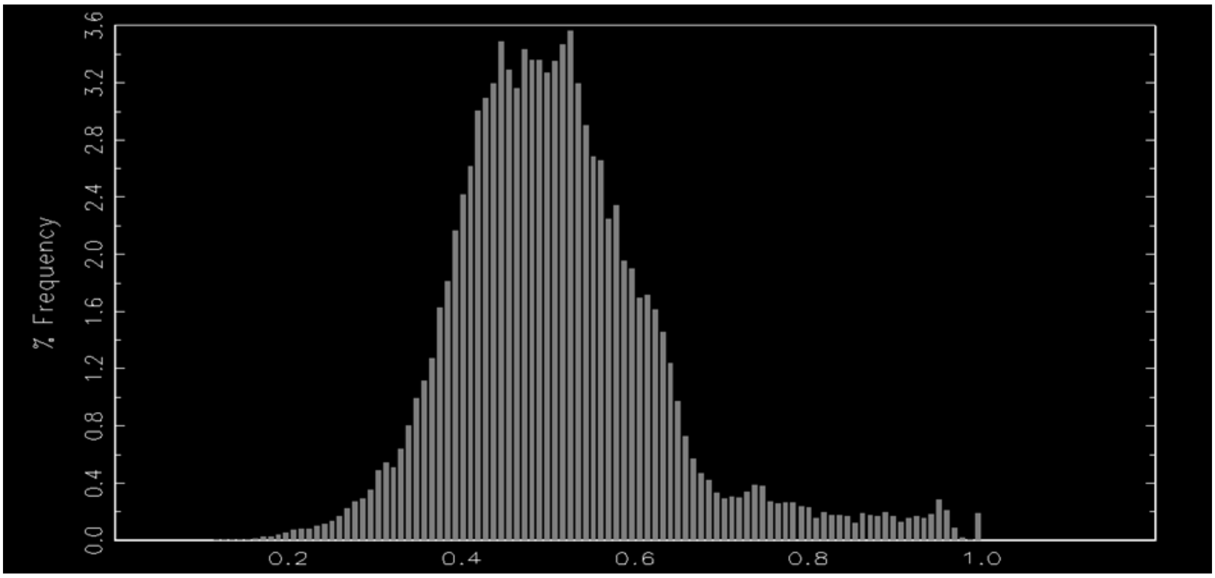


Figure C.4: Propensity score distribution (with mediators) among the controls