NOTE

WHY DO TOUGHER CASEWORKERS INCREASE EMPLOYMENT?
THE ROLE OF PROGRAM ASSIGNMENT AS A CAUSAL MECHANISM

Martin Huber, Michael Lechner, and Giovanni Mellace*

Abstract—Previous research found that less accommodating caseworkers are more successful in placing unemployed workers into employment. This paper explores the causal mechanisms behind this result using semiparametric mediation analysis. Analyzing rich linked job seeker–caseworker data for Switzerland, we find that the positive employment effects of less accommodating caseworkers are not driven by a particularly effective mix of labor market programs but, rather, by other dimensions of the counseling process, possibly including threats of sanctions and pressure to accept jobs.

I. Introduction

CASEWORKERS in employment offices usually have dual roles of job seeker counseling, and monitoring and may differ in performing these tasks. Behnecke, Fröhlich, and Lechner (2010, henceforth BFL) found for Switzerland that less cooperative caseworkers increase the reemployment chances of those who are unemployed. Reconsidering their linked job seeker–caseworker data, this paper decomposes the average treatment effect on the treated (ATET) of having a less cooperative caseworker into an indirect effect through the assignment of an active labor market program (ALMP) and a direct effect comprising all remaining causal mechanisms. That is, we consider the costly tool of ALMPs as an explicit mediator of caseworkers’ counseling style.1 Applying a sequential conditional independence assumption and propensity score matching, our findings suggest that the total effect is driven by channels other than program participation, which increase employment by initially roughly 1.5 percentage points, though the effect levels off over time. In contrast, the indirect path is never economically or statistically significant, such that the success of noncooperative caseworkers apparently does not come from more effective ALMPs.

The literature on direct and indirect effects or mediation analysis (see Baron & Kenny, 1986) has recently moved toward flexible modeling (see, e.g., Flores & Flores-Lagunes, 2009; Imai, Keele, & Yamamoto, 2010; Huber, 2014). As a methodological contribution, we propose a propensity score matching approach for estimating direct and indirect effects on the treated based on two matching steps using two different propensity scores: the conditional probabilities of treatment (a) given covariates and (b) given covariates and mediators. On the empirical side, we are the first to evaluate the causal mechanisms of caseworkers’ placement success in a mediation framework using linked survey-administrative data on caseworkers and job seekers.

The remainder of this paper is organized as follows. Section II discusses the econometric framework. Section III describes the data and selection issues. Section IV presents the results.

II. Econometric Framework

A. Potential Outcomes and Different Causal Effects

We aim at disentangling the effect of a binary treatment (D) on some outcome (Y) into a direct effect and an indirect effect operating through a possibly multidimensional mediator (M). We denote by \( Y(d) \) and \( M(d) \) the potential outcome and mediator states under treatment \( d \in \{0,1\} \) and by \( \Delta_{D=1} = E[Y(1) - Y(0)|D = 1] \) the ATET. To disentangle \( \Delta_{D=1} \), 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 effect (\( \theta_{D=1}(d) \))—the fraction of the total effect that is not attributed to the mediator—and the indirect effect (\( \delta_{D=1}(d) \))—the fraction attributed to the mediator, on the treated population:

\[
\theta_{D=1}(d) = E[Y(1, M(d)) - Y(0, M(d))|D = 1],
\]

\[
\delta_{D=1}(d) = E[Y(d, M(1)) - Y(d, M(0))|D = 1], \quad d \in \{0,1\}. \tag{1}
\]

Concerning the direct effect, Vansteelandt and VanderWeele (2012) argue that focusing on \( M(1) \) appears to be the natural reference for treated subjects when the choice of reference levels is a priori hard to justify, because \( M(1) \) corresponds to the actual choice of the treated:

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

Because the ATET is the sum of the direct and indirect effects defined on opposite treatment states,

\[
\Delta_{D=1} = E[Y(1, M(1)) - Y(0, M(1))|D = 1] + E[Y(0, M(1)) - Y(0, M(0))|D = 1] = \theta_{D=1}(1) + \delta_{D=1}(0),
\]

the indirect effect \( \delta_{D=1}(0) \) corresponds to the difference between the ATET and \( \theta_{D=1}(1) \):

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

Received for publication December 5, 2014. Revision accepted for publication February 5, 2016. Editor: Bryan S. Graham.

* Huber: University of Fribourg; Lechner: University of St. Gallen, CEPR and PSI, CESifo, IAB, and IZA; Mellace: University of Southern Denmark.

We have benefited from comments by the editor, Bryan Graham; Judea Pearl; seminar/conference participants at the Universities of Melbourne, Sydney, Otago, Copenhagen, Madrid (Carlos III), the EALE conference in Ljubljana, the workshop on Applied Microeconomics in Odense, the COMPIE conference in Rome, the IAAE meeting in Thessaloniki; as well as three anonymous referees. Financial support from the Swiss National Science Foundation, grant SNF 100018_137769 for the project “Identifying causal mechanisms in empirical economics” is gratefully acknowledged.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00632.

1 In 2003, training programs cost on average 110 CHF per person and day (plus additional expenses of 11 CHF for, traveling, for example), employment programs 98 CHF (plus 12 CHF). See Lalive, Zehnder, and Zweimüller (2006).

2 Robins and Greenland (1992) and Pearl (2001) use the denomination pure/total and natural direct and indirect effects, respectively.

© 2017 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology
doi:10.1162/REST_a_00632
In this paper, we therefore aim at identifying and estimating equations (2) and (3).³

B. Identifying Assumptions and Estimation

We impose sequential conditional independence of the treatment and the mediator:⁴

Assumption 1: \( \{Y(0,m), M(0)\} \perp \!\!\!\!\perp D | X = x \) for all \( m \) and \( x \) in the common support.

Assumption 1 states that the joint distribution of the potential outcomes (for any \( m \)) and mediators under nontreatment 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 observables.

Assumption 2: \( Y(0,m) \perp \!\!\!\!\perp M|X = x, D = d \) for all \( m, d, \) and \( x \) in the common support.

By assumption 2, the potential outcome under nontreatment and the observed mediator are independent conditional on the covariates and the treatment.⁵ This implies that \( Y(0,m) \) is conditionally independent of (the common-world state) \( M(0) \) whenever \( D = 0 \) and of (the cross-world state) \( M(1) \) whenever \( D = 1 \).

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

By assumption 3, there exists no combination of \( M, X \) that predicts treatment receipt with probability 1.

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

- \( E[Y(0,M(1)) | D = 1] = \frac{E[E[Y | M, X, D = 1]]}{E[X | D = 1]} \) (4)
- \( E[Y(0,M(0)) | D = 1] = \frac{E[E[Y | X, D = 1]]}{E[X | D = 1]} \) (5)

with \( E[C | B = b] \) denoting the expectation of \( C \) taken over the distribution of \( A \) conditional on \( B = b \). Thus, \( \theta_{D=1}(1) \) and \( \delta_{D=1}(0) \) are identified.

Proof. See online appendix A.

Our theorem 1 and its proof are similar to theorem 1 of Imai et al. (2010), albeit the latter focus on the total population, while we consider the treated. Instead of conditioning on \( M \) and \( X \) as in equations (4) and (5), the results of Rosenbaum and Rubin (1983) imply that identification is also obtained conditional on the propensity scores

³ The question asked by focusing on equations (2) and (3) is, “To which extent do the noncooperative caseworkers’ current ALMP strategies and other mechanisms contribute to the overall effect?” This is different from, “How effective are specific combinations of caseworker types and programs, given that policymakers may prescribe either?” which may be addressed by multiple or dynamic treatment evaluation (see, e.g., Imbens, 2000, and Lechner, 2009). Answering the first question explores the causal mechanisms underlying the status quo, while answering the second may help improve on the status quo.

⁴ Our identifying assumptions differ from those of Vansteelandt and VanderWeele (2012), who allow for confounding of \( M \) but impose stronger assumptions on the potential outcomes.

⁵ Our assumptions require that \( Y(0,M(1)) \) and \( Y(0,M(0)) \) take the same value whenever \( M(1) = M(0) \). This would be violated if the actual content of a formally identical ALMP differed across \( D \) conditional on \( X \).

### Table 1.—Survey Question on Cooperativeness of the Caseworker

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cooperation is very important; the wishes of the unemployed person should be satisfied.</td>
<td>☐</td>
</tr>
<tr>
<td>2. Cooperation is important, but placements and ALMP should sometimes be assigned or declined in spite of unemployed person’s wishes.</td>
<td>☑</td>
</tr>
<tr>
<td>3. Cooperation is less important; I should assign placements and ALMP independent of the wishes of the unemployed person</td>
<td>☐</td>
</tr>
</tbody>
</table>

Fifty-two percent of the caseworkers chose option 1, 39% chose option 2, and 9% chose option 3. Only very few caseworkers did not respond to this question. They are dropped from the analysis.

\[ p_{mx}(M,X) = \Pr(D = 1 | M, X) \] and \( p_{mx}(M) = \Pr(D = 1 | X) \), which balance the distributions of \( (M, X) \) and \( X \), respectively.⁶

\[ \theta_{D=1}(1) = E[Y | D = 1] - E_{p_{mx}(M,X) | D = 1} [E[Y | p_{mx}(M,X), D = 0]] \]

\[ \delta_{D=1}(0) = E_{p_{mx}(M,X) | D = 1} [E[Y | p_{mx}(M,X), D = 0]] - E_{p_{mx}(M,X) | D = 1} [E[Y | p_{mx}(M,X), D = 0]]. \]

We therefore propose estimating the direct and indirect effects by propensity score–based matching of nontreated observations to the treated using (a) the estimated \( p_{mx}(M,X) \) to obtain an estimate of \( \theta_{D=1}(1) \) and (b) the estimated \( p_{mx}(X) \) to estimate the ATET from which the direct effect is subtracted to estimate \( \delta_{D=1}(0) \).

We use distance-weighted radius matching on the propensity score with bias adjustment (see Lechner, Miquel, & Wunsch, 2011) and model \( p_{mx}(X) \) and \( p_{mx}(M,X) \) by flexible probit specifications. Concerning inference, we apply block bootstrapping and (a) resample caseworkers along with their job seekers to account for clustering, (b) compute the bootstrap t-statistics of the effects, and (c) estimate the p-value as the smoothed share of absolute bootstrap t-statistics that exceeds the absolute t-statistic in the original sample (see Racine & MacKinnon, 2007).

### III. Empirical Implementation

#### A. Study Sample and Definition of Treatment, Mediators, and Outcomes

The study sample comprises individuals who registered at Swiss employment offices in 2003. Our administrative data include, among others, nationality, qualification, education, experience, profession, and industry of last job; participation in ALMPs; employment histories; and regional characteristics. They were linked to a caseworker survey on aims, strategies, and processes of employment offices and caseworkers. The sample definition follows BFL (see appendix B.1) and ultimately contains 1,284 caseworkers and 100,120 unemployed.

The survey contained a question on the importance of cooperation with the job seeker (see table 1).⁷ As in BFL, the treatment \( (D) \) is 1 if

⁶ Specifically, \( X \perp [D | p_{mx}(X) = p_{mx}(x) \text{ and } (M,X)] \perp [D | p_{mx}(M,X) = p_{mx}(m,x)] \) for all \( m, x \) in the common support.

⁷ Without interactions between \( D \) and \( M \) in the outcome model, the indirect effect could alternatively be identified as the effect of \( D \) on \( M \) multiplied with the effect of \( M \) on \( Y \). This approach could incorporate posttreatment confounders influenced by \( D \) that affect both \( M \) and \( Y \). There thus exists a trade-off between model flexibility and allowing for posttreatment confounding; see, e.g., Imai and Yamamoto (2013).

⁸ According to the survey, noncooperative caseworkers consider control and sanctions, job assignments, and employment programs as more important and counseling meetings and temporary wage subsidies as less important. They also mentioned assigning ALMPs to exert pressure and control their clients’ availability for jobs.
the caseworker chose option 2 or 3 (is less cooperative) and 0 otherwise. The mediator ($M$) is the first participation in an ALMP within six months after starting the unemployment spell. We categorize ALMPs in six mutually exclusive groups: job search training, personality course, language skill training, computer training, vocational training, and participation in an employment program or internship. Together with non-participation, this entails seven possible mediator states. Estimation is based on six dummies for each ALMP being the first program in 2003. Table B.2.1 in appendix B.2 provides the frequencies of ALMPs across treatment states.

The outcome ($Y$) is a binary employment indicator on monthly bases. An individual is considered employed if she has deregistered at the employment office and the exit state is employment. As the mediator causally precedes the outcome, we consider only employment states assessed from month 8 after caseworker assignment (at least one month after program start). At least 29 outcome periods are available. At the latest, job seekers 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 outcome from the second half of 2004 until the end of in 2006. We therefore estimate the effects over a period of 1.5 to 3 years after caseworker assignment.

B. Selectivity of the Treatment and the Mediators

Our strategy requires observing all confounders of the relationship of the outcome with both the treatment or the mediator, or both (conditional on treatment). BFL argue that treatment selection depends on the types of hired caseworkers, the assignment of unemployed to caseworkers, and the development of the caseworkers’ attitudes as they gain work experience. As attitudes may be related to job placement success, we control for caseworker characteristics like age, gender, education, work experience, and own unemployment experience. We also control for rules assigning the unemployed to caseworkers, as stated in the survey. Furthermore, caseworkers may react differently to various types of unemployed and the labor market environment. Therefore, we include a range of job seeker characteristics like gender, qualification, previous occupation, labor market history, and a caseworker-provided employability rating, as well as local labor market conditions.

Nevertheless, some individual characteristics, like skills not reflected in qualifications, attitude, and motivation, are unobserved but arguably associated with observables. For instance, motivation should correlate with the labor market history and the employability rating (which should furthermore also capture unobserved skills). We therefore believe that conditional on observables, variation in the assignment of caseworker types is likely exogenous and, for instance, related to the current availability and work load of caseworkers.

Characteristics that have been identified as controls for caseworker cooperativeness are also expected to influence selection into the programs. As Gerfin and Lechner (2002) and many others argued, the assignment of ALMPs is likely driven by the job seeker’s socioeconomic background, previous occupation and labor market experience, employability, and regional characteristics. In addition, caseworker characteristics might affect the assignment strategy. Conditional on these factors, exogenous variation in ALMPs likely comes from differences in their availability across employment offices. Finally, it is worth noting that program assignment usually takes place early in the unemployment spell, so that time-varying confounding of the mediators due to changes in relevant factors during the unemployment spell is likely a minor issue. The fact that we use a (short) six-month window for program start is also in favor of this argument.

Table B.2.2 in appendix B.2 provides descriptive statistics for a range of confounders by treatment and program status. While there is limited selection with respect to caseworker rigor, perhaps with the exception of regional aspects, selection into the programs is much stronger and driven by several factors. (See also table B.3 of appendix B.3 for the propensity score specifications.) Appendix C contains tests for whether the propensity scores balance the characteristics of treated and nontreated in matching estimation (tables C.1 and C.2) and suggest that balancing works well.

IV. Results

Figure 1 contains the ATET as well as the direct and indirect effects among the treated on employment from month 8 to 36. The three lines represent the total, direct, and indirect effects, and superimposed symbols imply that these particular effects are (pointwise) significant at the indicated level. The results suggest that initially, less cooperative caseworkers significantly increase employment by roughly 1.5 percentage points. Over time, however, the ATET vanishes and is not statistically significant 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 rigor, although the indirect mechanism through ALMPs is never significant.

Our estimates therefore suggest that the (at least initially) higher job placement rates of less cooperative caseworkers are not due to...
better programs but to other dimensions that possibly include the threat of sanctions or pressure to go to job interviews. This suggests that policymakers should be interested not only in the effective provision of ALMPs but also in the analysis of other dimensions of case-workers’ counseling style, which can apparently make a difference.

11 We observe actual sanctioning days but do not find any economically or statistically significant effect of counseling style on sanctioning. Cooperative and noncooperative caseworkers could nevertheless differ in terms of the (unobserved) threat of sanctions, which may also affect employment.

REFERENCES


