Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies

Marco Caliendo\textsuperscript{a,b,c,d}, Robert Mahlstedt\textsuperscript{b,a}, Oscar A. Mitnik\textsuperscript{c,b}

\textsuperscript{a} University of Potsdam, Germany  
\textsuperscript{b} IZA, Bonn, Germany  
\textsuperscript{c} DIW, Berlin, Germany  
\textsuperscript{d} IAB, Nuremberg, Germany  
\textsuperscript{e} Inter-American Development Bank, United States

A R T I C L E   I N F O

JEL classification:  
C21  
D04  
J68

Keywords:  
Matching  
Unconfoundedness  
Unobservables  
Selection bias  
Personality traits  
Active labor market policy

A B S T R A C T

The main concern for many evaluation studies is that controlling for individuals’ observed characteristics may not be enough to obtain valid treatment effects. We exploit a unique dataset that contains a rich set of administrative information on individuals newly entering unemployment in Germany, as well as several usually unobserved characteristics like personality traits, attitudes, expectations, social networks and intergenerational information. This allows us to empirically assess the effect of including these usually unobserved variables on the propensity score distribution, the matching quality, and the treatment effects obtained using unconfoundedness-based estimators. Our findings indicate that these variables play a significant role for selection into active labor market programs (ALMP), but do not make a significant difference in estimating treatment effects on wages and employment prospects. This suggests that the usually unobserved variables we analyze are not a threat to the validity of the estimated treatment effects, if comprehensive control variables of the type usually used in modern ALMP evaluations (which include labor market histories) are available. Our results also suggest that rich administrative data may be good enough to draw policy conclusions on the effectiveness of ALMPs.

1. Introduction

Evaluating the causal effects on outcomes of an intervention or treatment has become the key empirical objective in many areas of Economics, Statistics, and other fields like Sociology, Political Science, Epidemiology, and Medicine. Among the most exhaustively studied interventions are active labor market policies (ALMP), both using experimental and nonexperimental methods. After the influential study by LaLonde (1986) raised concerns on the ability of nonexperimental methods to replicate the results of ALMP experiments, a very large literature developed analyzing methodological aspects related to ALMP evaluation, and nonexperimental methods in general.\textsuperscript{1} A key ever-present question that nonexperimental ALMP evaluations face is whether the data can account fully for all the factors that explain both the participation in, and the outcomes of, a program. The objective of this paper is to address this question, relying on unique data on several characteristics usually not observed in the context of ALMP evaluations, for individuals entering unemployment in Germany.

If the assignment to a program is non-random, assumptions are needed to identify the treatment effects of interest. One of the most popular approaches is based on the unconfoundedness or conditional independence assumption (Heckman et al., 1999; Imbens and

\textsuperscript{1} The authors thank the editor, Conny Wunsch, and two anonymous reviewers, as well as Michela Bia, Jose Galdo, John Ham, Rafael Novella, Hilmar Schneider and especially Jeeyong Sohn for helpful comments and suggestions. We also thank the participants in seminars and conferences at IAB Nuremberg, CEPS Luxembourg, the German Economic Association Meeting 2014, LIESR Luxembourg, the Impact Evaluation Network 8th Annual Meeting and the 5th Joint IZA/IFAU Conference on Labor Market Policy Evaluation for valuable comments. This study uses the IZA/IAB Linked Evaluation Dataset (for more information, see Arri et al., 2014, and Eberle et al., 2017). The IAB (Nuremberg) kindly gave us permission to use the administrative data. The linked administrative data were prepared and provided for this research project only. A supplementary Online Appendix is available at http://dx.doi.org/10.1016/j.labeco.2017.02.001. A previous version of this paper circulated as "Unobservable, but Unimportant? The Influence of Personality Traits (And Other Usually Unobserved Variables) for the Evaluation of Labor Market Policies". The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

\textsuperscript{2} Corresponding author at: University of Potsdam, Chair of Empirical Economics, August-Bebel-Str. 89, 14482 Potsdam, Germany.

\textsuperscript{3} E-mail addresses: caliendo@empwifo.uni-potsdam.de (M. Caliendo), mahlstedt@iza.org (R. Mahlstedt), oscar@mitnik.net (O.A. Mitnik).

\textsuperscript{4} See Heckman et al. (1999) for a survey of the ALMP evaluation literature and the early debate on the LaLonde (1986) study; see Khvie (2010) and Card et al. (2010) for an overview of ALMP evaluation in Europe; see Dehejia and Wahba (1999, 2002), Smith and Todd (2005) and Dehejia (2005) on the debate on using propensity score matching to evaluate the training program in the LaLonde study; see Imbens and Wooldridge (2009) for a recent survey of econometric methods used in program evaluation.

http://dx.doi.org/10.1016/j.labeco.2017.02.001

Received 20 January 2016; Received in revised form 15 December 2016; Accepted 15 February 2017

Available online 22 February 2017

0927-5371/ © 2017 Elsevier B.V. All rights reserved.
Wooldridge, 2009). In a binary setting where units are either treated or used as comparisons (controls), the assumption implies that after controlling for differences in observed covariates between the two groups, any remaining differences are as if they had been generated by random assignment to the groups. In the context of ALMP evaluations this implies that researchers need to observe all the variables that affect both treatment participation and labor market outcomes. The main concern is that the unconfoundedness assumption is not realistic in many cases, implying that there may be unobserved characteristics that simultaneously explain the particular treatment individuals received and the outcome of interest. In this case, estimators based on the unconfoundedness assumption – e.g. propensity score matching and weighting – become biased, either under- or overestimating the causal effects of the treatment.

Looking back at the last decade, the developments are twofold. On the one hand, many countries now offer access to (very) informative and complete administrative data – including detailed information on the labor market histories of individuals – increasing the likelihood that the unconfoundedness assumption is satisfied. On the other hand, the recent literature showing the influence of variables such as personality traits or preferences on economic outcomes (e.g. Heckman et al., 2006; Osborne Groves, 2005; Bowles et al., 2001), should be a cause of concern about the validity of the unconfoundedness assumption; these variables might be important on many dimensions in the context of ALMP (e.g. job search behavior, selection into programs, overall labor market performance) but have not been used previously as conditioning variables in this context.

In this paper we address this concern explicitly. We focus on a class of estimators that rely on comparing treated and control individual based on the propensity score and exploit a combination of rich administrative and survey data for a fresh inflow sample into unemployment in Germany. The data not only contain “typical” administrative-based information (similar to many other ALMP evaluations, particularly in Europe), but also information on characteristics usually not observed in the context of ALMP evaluations, like personality traits, attitudes, expectations, social networks and intergenerational information.3 This allows us to empirically assess how estimators based on the unconfoundedness assumption perform when alternatively including or not these usually unobserved variables. The key idea is that even if individuals in the treatment and control groups have similar values of their estimated propensity scores (based on the usually observed variables) they could still differ in the usually unobserved variables. Our paper relates to the prior literature dealing with the sensitivity of unconfoundedness-based estimators. Imbens (2003) and Ichino et al. (2008) have proposed methods to assess the sensitivity of unconfoundedness-based estimators to the presence of unobserved variables. With methodological differences in their approaches, these studies try to assess how large should the effect of hypothetically not observed variables be to invalidate the results obtained from applying propensity score-based estimators in different situations. Lechner and Wunsch (2013) explore, using a German dataset, how sensitive matching estimators are to the inclusion of a variety of usually observed (but rich) characteristics, and find that those rich characteristics can remove selection bias. Our paper also relates to the literature that tries to identify the bias from unobservables by using the amount of selection on observables (e.g. Altonji et al., 2005; Oster, 2016).

Building upon this previous literature, we estimate treatment selection models using alternative sets of variables, for three typical ALMP programs – short-term training, long-term training and wage subsidies. We examine the resulting propensity score distributions, ranks and matching quality. Based on these selection models we estimate average treatment effects on the treated, and compare the effects associated to the alternative variable sets. Our findings indicate that personality traits and other usually unobserved variables play a substantial role for selection into treatment. However, comprehensive control variables (including labor market histories) are able to operate as reasonable proxies for the information provided by the usually unobserved variables. Thus, the differences in treatment effects between including and excluding the usually unobserved variables are in general small. Although, our setting is similar to that of evaluation studies in many countries, it should be noted that evaluating other programs and using different sets of control variables or different evaluation approaches could lead to different conclusions. Nevertheless, our results indicate that the usually unobserved variables we analyze are not a threat to the validity of the treatment effects and suggest that rich administrative data that includes detailed labor market histories may be good enough to draw policy conclusions on the effectiveness of specific active labor market policies.

The paper is structured as follows. The next section gives a short summary on the identification of treatment effects and the role of potentially unobserved variables. Section 3 describes the institutional background and the dataset, and presents some descriptive statistics. Section 4 presents the results, while Section 5 concludes.

2. Usually unobserved variables and treatment effects

We base our discussion on the well known potential outcomes framework (Roy, 1951; Rubin, 1974) and focus on the usual parameter of interest in most evaluation studies, the average treatment effect on the treated (ATT):

\[ t_{ATT} = E(Y_i^1|D_i = 1) - E(Y_i^0|D_i = 1). \]  

(1)

\( Y_i^1 \) and \( Y_i^0 \) are potential outcomes for individual \( i \) with and without treatment and \( D_i \) is a treatment indicator (equal 1 if individual \( i \) received treatment). The last term on the right hand side of Eq. (1) is not observed and using the realized outcomes of non-participants instead, leads to a bias if participants and non-participants are selected groups who would have different potential outcomes even in the absence of treatment. To correct for this selection bias in nonexperimental studies, propensity score matching estimators rely on the conditional independence assumption (CIA), which implies that conditional on the propensity score \( P(X_i) = P(D_i = 1|X_i) \), where \( X_i \) is a set of observed characteristics, the counterfactual outcome is independent of the treatment. The CIA is a strong assumption and its justification depends crucially on the availability of data which allow the researcher to control for all relevant factors that simultaneously influence the participation decision and the potential outcomes. If there are unobserved variables which affect assignment into treatment and the potential outcomes simultaneously, a hidden bias might arise to which matching estimators are not robust (see, e.g. Rosenbaum, 2002, for an extensive discussion). Let us assume that the participation probability is determined by a set of variables \( W = (X, U) \), where the variables in \( X \) are observed, but the variables in \( U \) are not. Then the participation probability can be specified as:
دریافت فوری
متن کامل مقاله
امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات