Unobservable, but Unimportant? The Relevance of Usually Unobserved Variables for the Evaluation of Labor Market Policies

Marco Caliendo[*] **Robert Mahlstedt**[†]
Oscar A. Mitnik[‡]

forthcoming in: Labour Economics
February 15, 2017

Abstract

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 uncounfoundedness-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.

Keywords: Matching, Unconfoundedness, Unobservables, Selection Bias, Personality Traits, Active Labor Market Policy
JEL codes: C21, D04, J68

[*]University of Potsdam, IZA Bonn, DIW Berlin, IAB Nuremberg; e-mail: caliendo@empwifo.uni-potsdam.de. Corresponding address: University of Potsdam, Chair of Empirical Economics, August-Bebel-Str. 89, 14482 Potsdam, Germany. Tel: +49 331 977 3225. Fax: +49 331 977 3210.

[†]IZA Bonn, University of Potsdam; e-mail: mahlstedt@iza.org.

[‡]Inter-American Development Bank and IZA Bonn; e-mail: oscar@mitnik.net.

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 Jeffrey A. Smith 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, LISER 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 Arni, Caliendo, Künn, and Zimmermann, 2014). 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 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.

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.[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, LaLonde, and Smith, 1999; Imbens and 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.[2] 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

[1]See Heckman, LaLonde, and Smith (1999) for a survey of the ALMP evaluation literature and the early debate on the LaLonde (1986) study; see Kluve (2010) and Card, Kluve, and Weber (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.

[2]Even though the literature uses "selection on observables" as a way of referring to the unconfoundedness assumption, and the term "unobservables" is also commonly used, we prefer to use the term "unobserved" to highlight the fact that the observability of a particular variable will vary for different contexts and data.

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, Stixrud, and Urzua, 2006; Osborne Groves, 2005; Bowles, Gintis, and Osborne, 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 individuals 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, Mealli, and Nannicini (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, Elder, and Taber, 2005; Oster, 2014).

Building upon this previous literature, we estimate treatment selection models using alterna-

[3]For example Gerfin, Lechner, and Steiger (2005) for Switzerland, Sianesi (2004) for Sweden, and Lechner, Miquel, and Wunsch (2011) and Biewen, Fitzenberger, Osikominu, and Paul (2014) for Germany, use comprehensive administrative data in order to evaluate ALMP programs in (Western) European countries. However, those studies generally lack information about personality traits, attitudes and expectations.

tive 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 descriptives 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):

$$\tau_{ATT} = E(Y_i^1 \mid D_i = 1) - E(Y_i^0 \mid D_i = 1). \tag{1}$$

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 equation (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 non-experimental studies, propensity score matching estimators rely on the conditional independence assumption (CIA), which implies that conditional on the propensity score $P(X_i) =$

$Pr(D_i = 1 \mid X_i)$, where X_i is a set of observed characteristics, the counterfactual outcome is independent of treatment.[4] 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:

$$P_i = P(X_i, U_i) = P(D_i = 1 \mid X_i, U_i) = F(\beta X_i + \gamma U_i). \tag{2}$$

The study is free of any bias if γ is zero. Otherwise, two individuals with the same observed covariates X would differ in their odds of receiving treatment by a factor that involves the parameter γ and the difference in their unobserved covariates U.[5] These unobserved differences with respect to the chance of receiving the treatment would create a *hidden bias* when the outcome is correlated with U, after conditioning on X.

The extent to which γ and U play a role will depend on the empirical context. The importance of the unobserved characteristics U clearly depends on the extent of the observed characteristics. A more informative set of control variables X reduces the likelihood that, after controlling for X the resulting U has an effect on the participation decision. Previous studies suggest that socio-demographic and regional information as well as labor market histories of participants play an important role when evaluating treatment effects (e.g. Mueser, Troske, and Gorislavsky, 2007; Heckman, Ichimura, Smith, and Todd, 1998). Especially, the improved availability and quality of administrative data in recent years has allowed researchers to better understand the effects of certain characteristics on potential treatment effects in a systematic way (e.g. Lechner and Wunsch, 2013; Huber, Lechner, and Wunsch, 2013; Biewen, Fitzenberger, Osikominu, and Paul, 2014). However, at the same time a variety of studies shows the importance of variables previously not extensively considered in economics in general and for ALMP evaluations in particular, like personality traits (Nyhus and Pons, 2005), cognitive and non-cognitive skills (Heckman, Stixrud, and Urzua, 2006) or preferences and attitudes (Pannenberg, 2010; Belzil and Leonardi, 2007).

[4] In addition to the CIA, we also assume overlap which implies that there are no perfect predictors which determine participation, i.e. $Pr(X_i) < 1$, for all i.

[5] This can be easily seen if, as in Rosenbaum (2002), we assume we have a matched pair of individuals i and j and that F is the logistic distribution. The odds ratio that the individuals receive treatment is given by $\frac{\exp(\beta x_i + \gamma u_i)}{\exp(\beta x_j + \gamma u_j)}$. Successful matching implies that the X-vector cancels out, making the odds ratio equal to $exp[\gamma(u_i - u_j)]$.

In this context several variables which are usually not observed when evaluating labor market policies, might be of special interest. For example, Mueller and Plug (2006) find for the U.S. that the 'Big Five' personality traits – extraversion, agreeableness, conscientiousness, neuroticism and openness – have an impact on earnings similar to that found for cognitive abilities. Similar results are found for the Netherlands. Moreover, several empirical studies investigate how an individual's locus of control might be related to labor market performance. Locus of control refers to a general expectation about internal versus external control of reinforcement (Rotter, 1966). People with a more external locus of control believe that much of what happens in life is beyond their control, while people with an internal locus of control see life's outcomes as dependent on their own decisions and behavior. Several studies find a statistically significant effect of the locus of control on individual earnings (e.g. Andrisani, 1977; Heineck and Anger, 2010; Semykina and Linz, 2007) and job search strategies (e.g. Caliendo, Cobb-Clark, and Uhlendorff, 2014; McGee, 2015).

Another strand of the literature points out the importance of expectations in general (e.g. Spinnewijn, 2015), but also specifically about ALMP participation (e.g. Black, Smith, Berger, and Noel, 2003; van den Berg, Bergemann, and Caliendo, 2009), for the search behavior of unemployed job seekers. Moreover, the intergenerational transmission of human capital (Black, Devereux, and Salvanes, 2005) and attitudes (Dohmen, Falk, Huffman, and Sunde, 2012), as well as social networks (Montgomery, 1991; Bayer, Ross, and Topa, 2008) seem to play an important role in determining an individual's labor market performance. Finally, workers geographically mobility is also a driving factor of their economic outcomes (e.g. Yankow, 2003), while car access seems be of special importance (Gurley and Bruce, 2005). Combining these different strands of literature the natural question that arises is whether these variables also play a role when evaluating the effects of active labor market programs.

3 Institutional Background, Data and Descriptives

3.1 Institutional Background

Germany has a long tradition of ALMP and the German Social Security Code provides a large set of programs geared towards helping unemployed individuals, like training programs, wage subsidies, job creation schemes, start-up subsidies or benefits to increase the job-seeker's labor market mobility. Table 1 shows the entries into different programs in Germany between 2005 and 2011. As they are most relevant (in terms of number of participants) for supporting unemployed job-seekers and very typical for many OECD countries, we investigate the effect of three programs in detail: 1) Short-term training, 2) long-term training and 3) wage subsidies. While the short-term training represents a more recent group of programs shifting the focus towards more 'activating' elements, long-term training and wage subsidies represent more traditional programs, which aim to remove disadvantages in education, work experience or productivity. Since these programs represent very different reintegration strategies and are targeted at different types of unemployed individuals, this potentially allows us to examine the role of usually unobserved variables for these different selection processes. Let us briefly summarize some institutional details for these three programs.

[Insert Table 1 about here]

Short-term training measures, introduced in 1998, have a maximum duration of eight weeks. The courses can either serve as test of the participant's occupation-specific aptitude, or aim to improve the general employability. For example, the courses teach the unemployed how to apply effectively for a new job or how to behave in job interviews, but can also consist of computer or language classes. Some of the courses impart knowledge on starting a business to founders of start-ups, while others are concerned with the special needs of certain 'hard-to-place' job-seekers. Caseworkers can also use them to attain additional information on the participant's abilities and willingness to work. Courses are conducted full- or part-time and last from two days up to eight weeks; an individual's time spent in short-term training programs is limited to twelve weeks in total. While in a short-term training program an unemployed person cannot earn additional wages; however she continues to receive unemployment benefits and coverage of the costs associated to participation (e.g. transportation, child care, see Wolff and Jozwiak, 2007).

Long-term training programs have been a well established part of the German labor market policy for many decades. These programs can last from three months to up to three years. Historically, a caseworker would assign an unemployed individual to a specific course aimed at improving her occupational skills, and facilitating reintegration into the labor market. Previous studies find positive effects only in the very long-run (e.g. Fitzenberger, Osikominu, and Völter, 2008; Lechner, Miquel, and Wunsch, 2011) or even partly negative effects on employment (e.g. Lechner and Wunsch, 2008). With the 'Hartz reforms' at the beginning of the century, the German government reduced the usage of (long) vocational training programs. From 2003 onwards caseworkers no longer choose a specific course for the unemployed but hand out a training voucher to the job-seeker who is then allowed to find an appropriate training program for herself (see Bernhard and Kruppe, 2012; Doerr, Fitzenberger, Kruppe, Paul, and Strittmatter, 2014).

Wage subsidies are one of the oldest instruments used to reintegrate unemployed individuals into the labor market. The aim of the subsidy is to reduce the labor costs for the firm, potentially bridging any deficiencies in a worker's productivity. Wage subsidies (or temporary employment with a wage subsidy) can also be used as a screening device, lowering uncertainty and, hopefully, creating stable employer-employee relationships. Whether or not an unemployed person is supported with a targeted wage subsidy is a decision that is made by her caseworker. In addition, the caseworker determines the properties of the subsidy (restricted by the legal framework and guidelines): up to 50 percent of the monthly wage can be covered by the subsidy for at most 12 months. Extensions are possible if the wage subsidy aims at the integration of older or handicapped workers. Employers of subsidized workers agree to employ workers who are younger than 50 years for a follow-up period after the subsidy ends. This follow-up period is usually as long as the subsidized period itself. In case the worker is dismissed for reasons that are not attributable to her performance, the employer has to return a portion of the subsidy. Previous research indicates relatively large favorable effects on the employment prospects of hard-to-place workers using a matching approach (e.g. Bernhard and Wolff, 2008; Jaenichen and Stephan, 2011). However, Schünemann, Lechner, and Wunsch (2015) cast doubts on the findings of the prior literature due to methodological concerns. They argue that propensity score matching based on typically observed individual characteristics, including socio-demographic information and labor market histories, is unlikely to be sufficient for the evaluation of wage subsidies programs, since the receipt of the subsidy is conditional on being employed, which is not an exogenous factor. Exploiting a regression discontinuity design, their study does not find any significant impact of wage subsidies on job finding rates. This is a very important issue to which we will return when interpreting our findings.

3.2 The IZA Evaluation Dataset and our Estimation Sample

This study uses the *IZA/IAB Linked Evaluation Dataset* which combines survey information and administrative data on individuals who entered unemployment between June 2007 and May 2008 in Germany (see Caliendo et al., 2011). The dataset contains a 9% random sample, from the monthly unemployment inflows of approximately 206,000 individuals identified in the administrative records, who are selected for interview. From this gross sample of individuals aged between 16 and 54 years, representative samples of about 1,450 individuals are interviewed each month so that after one year twelve monthly cohorts were gathered (see Arni et al., 2014, for details on representativeness etc.). The first wave of interviews takes place shortly after the entry into unemployment. Besides the extensive set of individual-level characteristics and labor market outcomes, the individuals are asked a variety of non-standard questions about search behavior, social networks, psychological factors, cognitive and non-cognitive skills, subjective assessments on future outcomes, and attitudes. For the 88% of individuals who agreed, these survey data were then merged to administrative information from the *Integrated Employment Biographies* (IEB) provided by the Institute for Employment Research (IAB).[6] The IEB integrates different sources, e.g., employment history, benefit recipient history, training participation history and job search history and therefore contains detailed information on employment subject to social security contributions, unemployment and participation in active labor market policy including wages and transfer payments. The data additionally include a broad range of socio-economic characteristics including education, family status and health restrictions. The data do not contain information about working hours or periods in self-employment, working as a civil servant, or time spent in inactivity. Altogether, this amounts to a total of 15,274 realized interviews with a time lag from seven to fourteen weeks between the unemployment registration and the interview.

For the purpose of the study, we restrict our estimation sample to all individuals who are still unemployed and do not participate in any ALMP program when the interview takes place. We define job seekers as participants if they attend short- or long-term training, or receive a wage subsidy within the first twelve months after the entry into unemployment, and as non-participants if they do not participate in any ALMP program within this period. This leaves us with 4,934 non-participants, while 1,803 individuals participate in short-term training, 783 in long-term training and 542 receive a wage subsidy.[7]

[6]This study is based on a weakly anonymized sample of the Integrated Employment Biographies by the IAB (V.901).

[7]The choice of the period for the split is arbitrary and could be debated (see Sianesi, 2004); nevertheless it is a standard procedure in the evaluation of ALMP. In our case, choosing 12 months as the treatment period covers about 89% of all individuals who participate in an ALMP program within our complete observation period of 30 months and ensures that we observe individuals for a sufficiently long time window after the treatment. Moreover, increasing the treatment period, has the disadvantage that the non-participation in later periods is to some extent simply the consequence of a successful job search in earlier periods. Therefore, it becomes less clear whether the estimated effects are causal to the program participation. Alternatively, duration models would allow us to control for the exact timing of the treatment, however additional distributional assumptions would be necessary (see for example Card, Kluve, and Weber, 2010, for an overview of potential estimation strategies when evaluating 199 worldwide ALMP programs).

3.3 Some Descriptives for Outcomes and (Usually Un)Observed Variables

We observe every job-seeker in our sample for a period of 30 months after entering unemployment. To evaluate the influence of usually unobserved variables on the treatment effects we focus on labor market outcomes which are typically used in the evaluation of ALMP programs. In particular we concentrate our analysis on the employment probabilities at the end of our observation period after 30 months and the cumulative earnings within the observation period. The upper part of Table 2 shows the differences between participants and non-participants with respect to these labor market outcomes. We observe no (statistically) significant differences with respect to their employment probability between non-participants and, respectively, participants in short- and long-term training, while for recipients of a wage subsidy, the raw employment probabilities are higher after 30 months. However, the cumulative earnings are significantly lower for participants in short- and long-term training, but higher for recipients of wage subsidies.[8]

[Insert Table 2 about here]

Additionally, the lower part of Table 2 shows differences with respect to the main covariates of interest in our study – the *usually unobserved variables.* The first category, *personality traits*, include the 'Big Five' factors (except for agreeableness due to missing items) (see Digman, 1990, for an overview) and locus of control. It can be seen that participants in short-term training show a higher level of neuroticism, a lower internal locus of control and a lower level of openness. The latter is also true for participants in long-term training, while they additionally have a lower level of extraversion and are more conscientious. For recipients of wage subsidies, the only difference to non-participants can be observed with respect to their level of conscientiousness. Second, the *intergenerational variables* contain information on the father of the survey participant. It should be noted that for participants in short-term training and wage subsides, father's education is significantly lower and the father was more likely to be a blue-collar worker when the individuals were 15 years old. Moreover, we account for differences with respect to *social networks* and proxies for *labor market flexibility*, like being a car owner or having problems with childcare. Overall, there are only minor differences with respect to these variables. Individuals participating in short-term training have fewer good friends and participants in long-term training are less likely to have good contacts to their neighbors. Finally, we include variables measuring the individuals *life satisfaction* and *ALMP expectations.* All groups of participants report a lower

[8] It should be noted that the lower cumulative earnings of participants in training programs are induced by 'locking-in' effects, i.e. lower employment in early periods due to a reduction of participants' search effort before and during program participation (e.g. Van Ours, 2004; Jespersen, Munch, and Skipper, 2008).

life satisfaction and a higher ex ante probability of participating in a program.

When considering typical covariates, like socio-demographics and labor market histories, it should be noted that ALMP participants are in general more likely to be female, less likely to have health problems and spent more time in employment in the past. Except for long-term training, participants have also a lower level of education, they are less likely to hold an upper secondary school leaving or an university degree, and earn lower (daily) wages before entering unemployment. Moreover, recipients of wage subsidies have a higher probability of living in East-Germany, while participants in both training programs predominately live in West-Germany (compared to non-participants).

4 Empirical Results

4.1 Estimation Strategy

The objective of our study is to examine how estimators based on the unconfoundedness assumption perform when alternatively including – or not – usually unobserved variables. The implementation of propensity score matching and weighting estimators is a two-step-procedure where in a first step the participation model is estimated. The resulting participation probabilities are then used in a second step to match participants with similar non-participants. In the next subsections we evaluate the effect of the usually unobserved variables on each of these steps.

To conduct our analysis in a systematic way, we start by estimating different propensity score specifications for each ALMP program. Since labor market histories can be expected to play a key role in this context, we define two types of baseline specifications. The first only includes socio-demographic characteristics, household characteristics, local economic conditions and variables related to unemployment entry, while the second adds detailed information on short- and long-term labor market history-related variables. For each of the two baseline specifications, we then subsequently include each group of usually unobserved variables and evaluate their impact on the propensity score based on various measures.

Afterwards, we focus on three propensity score specifications and analyze implications on the propensity score distributions, the matching quality and the treatment effects in more detail. The choice of these models is motivated by their particular relevance in the context of ALMP evaluations. First, in the *standard* model we include all variables that have been typically used when evaluating these programs, including socio-demographic characteristics (and related variables), as well as short- and long-term labor market histories. This specification provides a reference model including variables which are consistently found to be key drivers of selection into training (e.g. Dolton and Smith, 2011; Lechner and Wunsch, 2013). Second, in the *auxiliary* model we explore the effects of replacing the labor market histories with the full set of usually unobserved variables. With this model we assess the extent to which the usually unobserved covariates U provide information that is similar (or not) to that provided by the labor market histories. This is of special interest, as U contains many characteristics that are typically assumed to be relatively stable over time, e.g., personality traits (see e.g. Heineck and Anger, 2010; Cobb-Clark and Schurer, 2013) and intergenerational variables, and can therefore be expected to influence the labor market outcomes of interest in similar ways as the employment and earning histories. Thus, the auxiliary model allows us to see whether both sets of variables can serve as a proxy for each other, and provides evidence for whether the usual claim in the ALMP evaluation literature that labor market histories can account for unobservables is actually justified. Finally, the *extended* model adds to the standard model the usually unobserved variables introduced in the auxiliary model and therefore exploits all available information. The key intuition behind our approach is that we identify the ATT imposing the assumption that the CIA holds alternatively for the usually observed covariates X (standard), or for the set X plus the set of usually unobserved covariates U (extended). These specifications are not, however, tests of the CIA assumption, nor is ours an exercise in model selection. Comparing the estimated treatment effects allows us to determine the sign and magnitude of potential hidden bias due the exclusion of the usually unobserved variables, but not whether the CIA holds for either model. A detailed depiction of the specifications and the full list of covariates is shown in Table A.1 in the Appendix.

4.2 Relevance for Propensity Score Estimation

We start the analysis by estimating the propensity score for each program using a logit model, as is standard in the literature. For the three treatments the control group contains only individuals which do not participate in any ALMP program within a period of 12 months after the entry into unemployment.[9]

As discussed in the previous section, we sequentially include the six groups of usually unobserved variables for two types of baseline specifications. First, only with socio-demographic and related variables and second, additionally including labor market histories. Finally, we jointly

[9]The alternative to estimating several reduced-form binary logit models would be estimating a multinomial logit model on the full set of potential treatment choices. However, we concentrate on binary choice models since these are more popular among typical evaluation studies. Moreover, Lechner (2002) shows that the results for matching estimators based on different models are relatively similar using Swiss data.

include all usually unobserved variables, which results in 16 specifications per treatment. For each model, three types of summary statistics are presented in Table 3:[10] i) *hitrate* indicates the share of correct predictions from the estimated model;[11] ii) *F-test* presents the p-values for *Wald* tests of the joint significance of groups of usually unobserved variables on the participation probability; and iii) *SSVY-test* presents p-values for the specification test (proposed by Shaikh, Simonsen, Vytlacil, and Yildiz, 2009) for whether the estimated propensity score models are misspecified.[12]

[Insert Table 3 about here]

Briefly summarizing the estimation results, it should be noted that ALMP expectations and life satisfaction have a statistically significant impact on the participation probability in all three programs. Unsurprisingly, a high expectation of ALMP participation, but also a low level of life satisfaction, increases the likelihood of receiving a treatment. Moreover, for both types of training programs, we find a significant effect of personality traits, while for short-term training also social networks seem to influence the participation decision. Among the personality traits, the results presented in the Appendix suggest that having an internal locus of control reduces the participation probability in short-term training, while a low level of extraversion increases the participation probability in long-term training.

When comparing the two different baseline specifications, it can be seen that for long-term training the model without labor market histories is more likely to be misspecified – as indicated by the low p-value associated to the SSVY-test – and that for this model the addition of the usually unobserved variables has a larger impact (8.8 vs. 4.9 percentage points) on the hitrate. The differences according to the choice of the baseline specification are less pronounced for short-term training and wage subsidies. Moreover, comparing the baseline specification of the second model (including socio-demographic characteristics and labor market histories) and the full model, it can be seen that including all usually unobserved variables increases the share of observations correctly predicted between 1.3 percentage points for wage subsidies and 4.9 percentage points for long-term training, while the major part of the increased hitrate, especially

[10] Full estimation results including marginal effects for each variable are presented in Tables A.2–A.4 in the Appendix.

[11] To calculate the hitrate we classify an observation as 1 if the estimated propensity score is larger than the sample average of individuals receiving the treatment (i.e. $\hat{P}(X) > \bar{P}$) and 0 otherwise (i.e. $\hat{P}(X) \leq \bar{P}$) (see Heckman and Smith, 1999; Caliendo and Kopeinig, 2008).

[12] We thank an anonymous reviewer for suggesting the use of this test, which is given by: $\hat{V}_n = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i} K\left(\frac{\hat{P}(X_i)-\hat{P}(X_j)}{h}\right) \hat{\varepsilon}_i \hat{\varepsilon}_j$, with kernel K, bandwidth h and $\hat{\varepsilon}_i = D_i - \hat{P}(X_i)$. We also thank Edward Vytlacil for sharing the Gauss code used by Shaikh, Simonsen, Vytlacil, and Yildiz (2009), which we adapted to Stata.

for long-term training, can be explained by the inclusion of ALMP expectations.

[Insert Figure 1 and Table 4 about here]

In the following, we focus on the three specific models (standard, auxiliary and extended) discussed in Section 4.1. First of all, it should be noted that for none of these three models we find evidence of misspecification. Moreover, Figure 1 presents the propensity score distributions for each specification, separately for the three types of treatments and by participation status. The figure shows the importance of the usually unobserved variables in explaining program participation: the distributions are in general affected by the introduction of these variables, except for non-participants when used as comparison group for the long-term training and wage subsidies treatment.

To further explore the relationship between the three propensity score specifications, Table 4 provides several correlation measures between them. When comparing the extended and the standard specification, the correlation between propensity scores, as well as between corresponding ranks is very high and similar for all programs, while much lower when comparing the standard and the auxiliary specifications. Additionally, we also consider the absolute difference in propensity scores in the extended (standard) model for individuals who have been matched in the standard (auxiliary) model. It can be seen that the average difference in the extended model is more than twice as large for short- and long-term training than for wage subsidies. This highlights the importance of the usually unobserved variables for propensity score estimation, especially for the two training programs. Finally, we employ a Wilcoxon signed-ranks test which compares the rank of the individual propensity score differences for the two specifications of interest. In contrast to a simple sign test, it accounts not only for the direction but also for the size of the differences (e.g. Wilcoxon, 1945). When comparing the standard and auxiliary specification for long-term training, it can be seen that including the usually unobserved variables instead of the labor market history shifts the propensity score distribution to the right for participants, and respectively to the left for non-participants, while it is exactly the other way around for wage subsidies.

It is useful to relate the characteristics of the different programs to their selection process. For all programs the results related to the usually unobserved variables are intuitive: training programs, especially long-term training, require a high degree of commitment by the trainees, to endure a program that can be as long as three years. It is not surprising then that personality traits matter for selection into training programs (as indicated by their joint significance). Wage subsidies aim to help individuals find employment in markets where the demand for labor may be weak for the particular skills of these individuals. Therefore, it is not unexpected that the labor market histories of the individuals have the main explanatory power in the selection process.

4.3 Consequences for Matching Quality

An important indicator of the quality of the matching, and of the propensity score specification, is the balancing in the distribution of covariates between participants and non-participants. One suitable indicator for balancing is the mean standardized bias (Rosenbaum and Rubin, 1983), which assesses the distance of the covariates before and after matching.[13] In Table 5 we present the mean standardized bias (MSB) for groups of variables and overall, for the different specifications. In our setting, it is especially interesting to assess the matching quality for the usually unobserved variables under the standard specification, which does not include these variables; it is a useful way to summarize the degree to which matching only on the socio-demographic characteristics and labor market history can proxy for matching on the usually unobserved variables.

[Insert Table 5 about here]

The first column of Table 5 presents the *raw* MSB, i.e. prior to matching, while the next three columns present the MSB when matching with the alternative specifications of the propensity score. Using the extended propensity score specification reduces the overall MSB down to 2.1 for short-term training, to 3.0 for long-term training, and to 3.5 for wage subsidies, all very low values. However, on closer examination we find substantial differences with respect to the different programs and groups of control variables. For all types of programs, we find the largest raw bias with respect to ALMP expectations (MSB ranging from 17.2 for wage subsides up to 47.3 for long-term training), while, except for life satisfaction, it is on a moderate level for the other groups of usually unobserved variables. Overall, the standard specification is not very successful in reducing the MSB for ALMP expectations (for all programs) and life satisfaction (especially for short- and long-term training). When considering the auxiliary model it can be seen that, especially for wage subsides, conditioning on the usually unobserved variables does not reduce the mean bias for labor market histories which indicates that both groups of variables can not be considered as good proxies for each other. These results are in line with our previous findings which have shown that expectations and measures for life satisfaction, but

[13]For each covariate X, it is defined as: $SB(x) = 100(\bar{x}_c - \bar{x}_t)/\sqrt{\frac{1}{2}(s^2_{xc} + s^2_{xt})}$ with $\bar{x}_c$ being the mean of the control group, $\bar{x}_t$ the mean of the treatment group, s^2_{xc} the variance of the control group and s^2_{xt} the variance of the treatment group.

also labor market histories, have a strong impact on the selection into all types of programs. More importantly, estimating the propensity score using only the standard variables does not appear to be successful in eliminating the differences in the usually unobserved variables, which appear as very important for the selection into treatment process.

4.4 Consequences for Treatment Effects

In this section we present the consequences of using the alternative propensity score specifications for the estimation of the treatment effects of each program. There are several possible estimators for the Average Treatment on the Treated (ATT) parameters we are interested in obtaining (e.g. Imbens and Wooldridge, 2009). For the sake of clarity, we focus our analysis on a particular estimator, kernel matching, which is heavily used in evaluation studies. When relying on kernel matching estimators researchers need to specify a kernel function and a bandwidth parameter.[14] We specify an Epanechnikov kernel, and a bandwidth of 0.06. In the Appendix we conduct a sensitivity analysis where we specify alternative estimators (inverse probability weighting, IPW, nearest neighbor and radius matching), and bandwidth parameters ($bw = 0.02$; $bw = 0.2$) for the kernel estimators. The estimation results are qualitatively similar for all types of estimators.

Table 6 presents the differences in mean outcomes (raw gap) as well as the ATT from using the kernel estimator. We use the same three specifications for the propensity score discussed above. As outcomes of interest we analyze the employment probability 30 months after the entry into unemployment and the cumulated earnings within 30 months. The left panel of the Table shows the ATTs, while the right panel calculates the difference in ATTs.

Shortly summarizing the estimated effects of the different programs we find no effect of short- and long-term training on the employment probability after 30 months, and we find negative effects of those programs on unconditional earnings. Regarding wage subsidies we find a positive and significant effect on the employment probability after 30 months and on cumulated earnings, although these positive results may suffer from an upward bias, as we discuss further below. Our estimation results on long-term training are in line with previous studies, that find negative effects on employment probabilities in the short-run (e.g Lechner and Wunsch, 2008; Doerr, Fitzenberger, Kruppe, Paul, and Strittmatter, 2014) (which most likely shows up as a negative effect on unconditional earnings). Also for wage subsidies we find similar positive effects, at least after 30 months, as Bernhard and Wolff (2008) or Jaenichen and Stephan (2011). With

[14] In contrast to the choice of the bandwidth parameter, where a trade-off between a small variance and an unbiased estimate of the true conditional mean function arises, the choice of the kernel type appears to be relatively less important in practice (see the discussion in Caliendo and Kopeinig, 2008; Galdo, Smith, and Black, 2008).

respect to short-term training, Biewen, Fitzenberger, Osikominu, and Paul (2014) report a short locking-in period after the treatment which is in line with the negative effect on cumulated earnings. However, in the long-run we find no positive effect on employment probabilities (e.g. Wolff and Jozwiak, 2007) nor a reduction of the unemployment duration (e.g. Hujer, Thomsen, and Zeiss, 2006) as found by prior studies, for different time periods and target groups, in this literature.

Our main interest is in the comparison of the estimated treatment effects using alternative propensity score specifications. First of all, it should be noted that for all outcome variables and treatments the estimated ATTs conditioning on the standard set of covariates differ significantly from the unconditional raw differences. However, when comparing the propensity score specifications, it can be seen that the overall differences across specifications are relatively small. When comparing the standard with the auxiliary specification, we only find statistically significant (small) differences for long-term training in one of the two outcomes variables, suggesting in that case that the labor market histories cannot serve completely as proxies for the usually unobserved variables. However, except for employment 30 months after entry for wage subsidies participants, we do not find a statistically significant – or an economically relevant – difference between the standard and the extended specifications for the different programs and outcomes. This supports the idea that a large part of the usually unobserved characteristics, especially those that are constant over time, can be captured by controlling for the prior labor market performance.

[Insert Table 6 about here]

The results for the ATTs seem surprising given the importance for the selection process that the usually unobserved variables seem to have, as indicated by the propensity score marginal effects and distributions, and the measures of matching quality.[15] It is clear that the distinct selection processes suggested by the propensity score specifications are not reflected in differences in the estimated treatment effects. Overall, our results make it clear that the variables in the standard model are able to capture most of the information contained in the usually unobserved variables which are relevant for the labor market outcomes we analyze. This point is probably even stronger if some of the usually unobserved variables are stable over time as, for example, personality traits are expected to be. In general, it is reasonable to expect that the higher is the

[15]Even though our methodological approach is not based on linear regression, our results are consistent with those of Oster (2014) who shows that the omission of unobserved variables, in the context of linear models, does not necessarily generate large changes in the coefficients associated to a treatment, even if their omission generates omitted variable bias.

correlation between the usually unobserved variables and the labor market history, the smaller the additional value of the usually unobserved variables will be.

Regarding wage subsidies it is important to reiterate that, as discussed by Schünemann, Lechner, and Wunsch (2015), the selection process into the program is expected to be more complex since participation additionally requires the presence of an employer who hires the job seeker. Therefore, propensity score matching based on typical individual characteristics might not be sufficient to model the selection process and hence the standard model would tend to overestimate the actual treatment effect. Our findings indeed show a small reduction of the estimated ATT when accounting for usually unobserved variables in the extended model. However, it should be noted that, in contrast to Schünemann, Lechner, and Wunsch (2015) who use a regression discontinuity design, we still find a positive effect which suggests that additional information, e.g. employer characteristics, might be necessary in order to obtain a valid treatment effect for wage subsidies.

5 Conclusion

The aim of this study was to investigate the effect of usually unobserved variables, like personality traits, attitudes, expectations, social networks and intergenerational information, on the selection into active labor market policy programs and on the estimated average treatment effects. The results present a clear picture. The usually unobserved variables matter in terms of the selection process into treatment, in a different manner for the individuals treated under each of the programs. This is consistent with the three programs representing distinct reintegration strategies targeted to different types of unemployed individuals. Even though we find that the usually unobserved variables matter for selection, when estimating the effects of ALMP programs on labor market outcomes in a second step, the overall influence of including or excluding them is rather small.

The relatively small overall impact on treatment effects of the usually unobserved variables seems to be explained by the comprehensive control variables in the standard propensity score specifications, including labor market history variables. Assuming that the usually unobserved variables are constant over time, they not only affect selection into programs and future labor market outcomes, but they are probably correlated with past labor market performance. Thus, conditioning on labor market histories implicitly captures a large part of the information contained in the usually unobserved variables. Our results show that given our set of usually unobserved characteristics, the influence of these variables on the effect of ALMP programs in Germany is generally limited when informative administrative data are available. This suggests that lacking these usually unobserved characteristics does not affect in a fundamental way the assessment of public policies: As long as a large enough set of covariates is available any expected biases associated to not observing some of the personality traits, expectations and socio-cultural characteristics, are likely to be sufficiently small as to not fundamentally affect policy conclusions.

Moreover, it is necessary to be prudent in generalizing our results outside the setting of this study: The effects can clearly differ among different types of programs, different countries and populations of interest, as well as for other types of unobserved variables. Nevertheless, our study does show that valid concerns about the role of unobserved variables, when using a "selection on observables" assumption for the estimation of treatment effects, may be less relevant when observable information is available that is sufficiently correlated to the unobservable variables. This clearly seems to be the case in settings, like in many European countries these days, where policy evaluation is based on detailed administrative data. Finally, our paper provides a sort of "road-map" for researchers interested in systematically assessing the sensitivity of their results to the inclusion of alternative sets of control variables, which can be used in the context of any observational study relying on the unconfoundedness assumption.

References

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," Journal of Political Economy, 113(1), 151–184.

Andrisani, P. J. (1977): "Internal-external attitudes, personal initiative, and the labor market experience of black and white men," Journal of Human Resources, pp. 308–328.

Arni, P., M. Caliendo, S. Künn, and K. Zimmermann (2014): "The IZA Evaluation Dataset Survey: A Scientific Use File," IZA Journal of European Labor Studies, 3:6.

Bayer, P., S. L. Ross, and G. Topa (2008): "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," Journal of Political Economy, 116(6), 1150–1196.

Belzil, C., and M. Leonardi (2007): "Can risk aversion explain schooling attainments? Evidence from Italy," Labour Economics, 14(6), 957–970.

Bernhard, S., and T. Kruppe (2012): "Effectiveness of Further Vocational Training in Germany–Empirical Findings for Persons Receiving Means-tested Unemployment Benefits," Schmollers Jahrbuch: Journal of Applied Social Science Studies/Zeitschrift für Wirtschafts-und Sozialwissenschaften, 132(4), 501–526.

Bernhard, S., and J. Wolff (2008): "Contracting out placement services in Germany: Is assignment to private providers effective for needy job-seekers?," Discussion Paper 05/2008, IAB, Nuremberg.

Biewen, M., B. Fitzenberger, A. Osikominu, and M. Paul (2014): "The Effectiveness of Public-Sponsored Training Revisited: The Importance of Data and Methodological Choices," Journal of Labor Economics, 32(4), 837–897.

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

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005): "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," American Economic Review, 95(1), 437–449.

Bowles, S., H. Gintis, and M. Osborne (2001): "Incentive-enhancing preferences: Personality, behavior, and earnings," American Economic Review, pp. 155–158.

Caliendo, M., D. Cobb-Clark, and A. Uhlendorff (2014): "Locus of control and job search strategies," Review of Economics and Statistics, 97(1), 88–103.

Caliendo, M., A. Falk, L. Kaiser, H. Schneider, A. Uhlendorff, G. van den Berg, and K. Zimmermann (2011): "The IZA Evaluation Dataset: towards evidence-based labor policy making," International Journal of Manpower, 32, 731–752.

Caliendo, M., and S. Kopeinig (2008): "Some practical guidance for the implementation of propensity score matching," Journal of Economic Surveys, 22(1), 31–72.

Cameron, A. C., and P. K. Trivedi (2010): Microeconometrics using stata. Stata Press College Station, TX.

Card, D., J. Kluve, and A. Weber (2010): "Active Labour Market Policy Evaluations: A Meta-Analysis*," The Economic Journal, 120(548), F452–F477.

Cobb-Clark, D. A., and S. Schurer (2013): "Two economists' musings on the stability of locus of control," The Economic Journal, 123(570), F358–F400.

DEHEJIA, R. (2005): "Practical propensity score matching: a reply to Smith and Todd," Journal of Econometrics, 125(1), 355–364.

DEHEJIA, R. H., AND S. WAHBA (1999): "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," Journal of the American Statistical Association, 94(448), 1053–1062.

DEHEJIA, R. H., AND S. WAHBA (2002): "Propensity Score-Matching Methods for Nonexperimental Causal Studies," Review of Economics and Statistics, 84(1), 151–161.

DIGMAN, J. M. (1990): "Personality structure: Emergence of the five-factor model," Annual Review of Psychology, 41(1), 417–440.

DOERR, A., B. FITZENBERGER, T. KRUPPE, M. PAUL, AND A. STRITTMATTER (2014): "Employment and Earnings Effects of Awarding Training Vouchers in Germany," IZA Discussion Paper No. 8454.

DOHMEN, T., A. FALK, D. HUFFMAN, AND U. SUNDE (2012): "The intergenerational transmission of risk and trust attitudes," Review of Economic Studies, 79(2), 645–677.

DOLTON, P. J., AND J. A. SMITH (2011): "The Impact of the UK New Deal for Lone Parents on Benefit Receipt," IZA Discussion Paper No. 5491.

FITZENBERGER, B., A. OSIKOMINU, AND R. VÖLTER (2008): "Get training or wait? Long-run employment effects of training programs for the unemployed in West Germany," Annales d'Economie et de Statistique, pp. 321–355.

GALDO, J., J. A. SMITH, AND D. BLACK (2008): "Bandwidth Selection and the Estimation of Treatment Effects with Unbalanced Data," Annals of Economics and Statistics / Annales d'Economie et de Statistique, 91/92, 189–216.

GERFIN, M., M. LECHNER, AND H. STEIGER (2005): "Does subsidised temporary employment get the unemployed back to work? Aneconometric analysis of two different schemes," Labour Economics, 12(6), 807–835.

GURLEY, T., AND D. BRUCE (2005): "The effects of car access on employment outcomes for welfare recipients," Journal of Urban Economics, 58(2), 250–272.

HECKMAN, J., H. ICHIMURA, J. SMITH, AND P. TODD (1998): "Characterizing selection bias using experimental data," Econometrica, 66(5), 1017–1098.

HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH (1999): "The economics and econometrics of active labor market programs," Handbook of Labor Economics, 3, 1865–2097.

HECKMAN, J. J., AND J. A. SMITH (1999): "The Pre-programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies," The Economic Journal, 109(457), 313–348.

HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," Journal of Labor Economics, 24(3), 411–482.

HEINECK, G., AND S. ANGER (2010): "The returns to cognitive abilities and personality traits in Germany," Labour Economics, 17(3), 535–546.

HUBER, M., M. LECHNER, AND A. STEINMAYR (2015): "Radius matching on the propensity score with bias adjustment: tuning parameters and finite sample behaviour," Empirical Economics, 49(1), 1–31.

HUBER, M., M. LECHNER, AND C. WUNSCH (2013): "The performance of estimators based on the propensity score," Journal of Econometrics, 175(1), 1–21.

Hujer, R., S. L. Thomsen, and C. Zeiss (2006): "The effects of short-term training measures on the individual unemployment duration in West Germany," Discussion paper, ZEW Discussion Papers.

Ichino, A., F. Mealli, and T. Nannicini (2008): "From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity?," Journal of Applied Econometrics, 23(3), 305–327.

Imbens, G. W. (2003): "Sensitivity to exogeneity assumptions in program evaluation," American Economic Review, 93(2), 126–132.

Imbens, G. W., and J. M. Wooldridge (2009): "Recent Developments in the Econometrics of Program Evaluation," Journal of Economic Literature, 47(1), 5–86.

Jaenichen, U., and G. Stephan (2011): "The effectiveness of targeted wage subsidies for hard-to-place workers," Applied Economics, 43(10), 1209–1225.

Jespersen, S. T., J. R. Munch, and L. Skipper (2008): "Costs and benefits of Danish active labour market programmes," Labour Economics, 15(5), 859–884.

Kluve, J. (2010): "The effectiveness of European active labor market programs," Labour Economics, 17(6), 904–918.

LaLonde, R. J. (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," American Economic Review, 76(4), 604–620.

Lechner, M. (2002): "Program heterogeneity and propensity score matching: An application to the evaluation of active labor market policies," Review of Economics and Statistics, 84(2), 205–220.

Lechner, M., R. Miquel, and C. Wunsch (2011): "Long-Run effects of public sector Sponsored Training in West Germany," Journal of the European Economic Association, 9(4), 742–784.

Lechner, M., and C. Wunsch (2008): "What did all the money do? On the general ineffectiveness of recent West German labour market programmes," Kyklos, 61(1), 134–174.

——— (2013): "Sensitivity of matching-based program evaluations to the availability of control variables," Labour Economics, 21, 111–121.

McGee, A. D. (2015): "How the Perception of Control Influences Unemployed Job Search," Industrial and Labor Relations Review, 68(1), 184–211.

Montgomery, J. D. (1991): "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," American Economic Review, 81(5), 1407–18.

Mueller, G., and E. Plug (2006): "Estimating the effect of personality on male and female earnings," Industrial and Labor Relations Review, pp. 3–22.

Mueser, P. R., K. R. Troske, and A. Gorislavsky (2007): "Using state administrative data to measure program performance," Review of Economics and Statistics, 89(4), 761–783.

Nyhus, E. K., and E. Pons (2005): "The effects of personality on earnings," Journal of Economic Psychology, 26(3), 363–384.

Osborne Groves, M. (2005): "How important is your personality? Labor market returns to personality for women in the US and UK," Journal of Economic Psychology, 26(6), 827–841.

Oster, E. (2014): "Unobservable Selection and Coefficient Stability: Theory and Evidence," University of Chicago Booth School of Business, mimeo.

Pannenberg, M. (2010): "Risk attitudes and reservation wages of unemployed workers: evidence from panel data," Economics Letters, 106(3), 223–226.

Rosenbaum, P., and D. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," Biometrika, 70(1), 41–50.

Rosenbaum, P. R. (2002): Observational Studies. Springer, New York.

Rotter, J. B. (1966): "Generalized expectancies for internal versus external control of reinforcement.," Psychological monographs: General and applied, 80(1), 1.

Roy, A. (1951): "Some Thoughts on the Distribution of Earnings," Oxford Economic Papers, 3(2), 135–145.

Rubin, D. (1974): "Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies," Journal of Educational Psychology, 66, 688–701.

Schünemann, B., M. Lechner, and C. Wunsch (2015): "Do long-term unemployed workers benefit from targeted wage subsidies?," German Economic Review, 16(1), 43–64.

Semykina, A., and S. J. Linz (2007): "Gender differences in personality and earnings: Evidence from Russia," Journal of Economic Psychology, 28(3), 387–410.

Shaikh, A. M., M. Simonsen, E. J. Vytlacil, and N. Yildiz (2009): "A specification test for the propensity score using its distribution conditional on participation," Journal of Econometrics, 151(1), 33–46.

Sianesi, B. (2004): "An evaluation of the Swedish system of active labor market programs in the 1990s," Review of Economics and Statistics, 86(1), 133–155.

Smith, J. A., and P. E. Todd (2005): "Does matching overcome LaLonde's critique of non-experimental estimators?," Journal of Econometrics, 125(1), 305–353.

Spinnewijn, J. (2015): "Unemployed but optimistic: Optimal insurance design with biased beliefs," Journal of the European Economic Association, 13(1), 130–167.

van den Berg, G., A. Bergemann, and M. Caliendo (2009): "The Effect of Active Labor Market Programs on Not-Yet Treated Unemployed Individuals," Journal of the European Economic Association, 7(2-3), 606–616.

Van Ours, J. C. (2004): "The locking-in effect of subsidized jobs," Journal of Comparative Economics, 32(1), 37–55.

Wilcoxon, F. (1945): "Individual comparisons by ranking methods," Biometrics bulletin, pp. 80–83.

Wolff, J., and E. Jozwiak (2007): "Does short-term training activate means-tested unemployment benefit recipients in Germany?," Discussion Paper 29/2007, IAB, Nuremberg.

Yankow, J. J. (2003): "Migration, job change, and wage growth: a new perspective on the pecuniary return to geographic mobility," Journal of Regional Science, 43(3), 483–516.

Tables and Figures

Table 1: Entries in ALMP Programs in Germany (in 1,000)

	2005	2007	2009	2011
Entries into unemployment	8,427	8,155	9,253	8,218
Entries into ALMP programs				
Short-term training	901	1,092	1,194	1,201
Long-term training	132	365	618	305
Wage subsidies	134	266	266	187
Job creation schemes	78	70	11	1
Start-up subsidies	91	126	137	137
Mobility assistances[a)]	211	352	21	—

Source: Statistic of the German Federal Employment Agency
[a)] Not separately accounted after 2009.

Table 2: Selected Descriptive Statistics by Treatment Status

	Non-participants	Short-term training	Long-term training	Wage subsidies
No. of observations	4,934	1,803	783	542
Labor market outcomes				
Regular employed 30 months after entry (t+30)	0.55	0.56	0.56	***0.68***
Cumulated earnings in € up to t+30	24,490	***22,075***	***20,615***	***29,273***
Usually unobserved variables				
A. Personality traits[a)]				
Openness	5.06	*5.00*	*4.97*	5.10
Conscientiousness	6.23	6.25	*6.30*	**6.34**
Extraversion	5.20	5.15	***5.07***	5.19
Neuroticism	3.77	*3.83*	3.78	3.71
Locus of Control	5.05	***4.95***	*5.00*	5.00
B. Intergenerational variables				
Father has upper sec. school leaving degree	0.16	***0.11***	0.15	**0.12**
Father was employed when person aged 15	0.84	0.83	0.85	0.86
Father's current age				
≤60 years	0.37	0.36	***0.29***	**0.31**
>60 years	0.34	0.33	***0.41***	0.33
Father was blue-collar worker when person aged 15	0.33	**0.36**	0.35	**0.38**
C. Social network				
Number of good friends outside family: less than two	0.09	***0.12***	0.11	0.09
Contacts to neighbors:[b)] good (1-3)	0.73	0.72	*0.70*	0.74
D. Labor market flexibility				
Childcare situation:[b)] bad (4-6)	0.06	0.06	0.07	0.05
Car ownership	0.60	0.61	**0.65**	***0.67***
E. Life satisfaction:[c)]				
Low (0-3)	0.10	0.11	0.10	*0.12*
High (7-10)	0.57	***0.52***	***0.50***	***0.48***
F. Expected ALMP probability:[c)]				
Low (0-3)	0.45	***0.30***	***0.25***	***0.37***
High (7-10)	0.31	***0.43***	***0.55***	***0.39***
Socio-demographic characteristics				
Female	0.46	***0.51***	***0.53***	0.47
Age in years	34.99	34.89	***36.88***	***37.75***
Married (or cohabiting)	0.38	0.39	*0.42*	0.40
German citizenship	0.95	0.94	0.95	0.95
Living West-Germany	0.67	**0.70**	**0.70**	***0.61***
Two children or more	0.13	0.13	***0.17***	0.14
Upper sec. school leaving degree	0.27	***0.18***	0.28	***0.22***
University degree	0.20	***0.13***	0.21	*0.17*
Health restriction or disability	0.08	**0.06**	*0.06*	***0.04***
Labor market history				
Employment status before UE: regular employed	0.64	0.65	**0.68**	***0.70***
Last daily wage in €	46.04	**43.51**	*48.71*	44.79
Last job was full-time employment	0.96	**0.94**	0.95	0.95
Months in regular employment				
in last 6 months	4.07	**4.21**	4.04	**4.32**
in last 2 years	15.02	***16.12***	*15.59*	***16.33***
in last 10 years	49.28	***52.43***	**52.95**	***54.98***
No. of employers				
in last 2 years	1.56	1.56	1.57	**1.65**
in last 10 years	2.90	2.94	2.99	***3.30***

Notes: All numbers denote shares unless otherwise indicated, measured at the entry into unemployment. Italic/bold/italic+bold numbers indicate statistically significant differences between each group of participants and non-participants at the 10/5/1%-level based on a two-tailed t-test on equal means.

Additional socio-demographics: Migration background, searching for full- or part-time employment, employment status of partner, local unemployment rate, month of entry into unemployment, time between entry into unemployment and interview.

Additional labor market history: Laid off by last employer, time with last employer, duration of last unemployment spell, months in unemployment/ ALMP program/ out of labor force in last 6 months/ 2 years/ 10 years, number of program participations/ unemployment spells/ out of labor force spells in last 2 years/ 10 years.

[a)] Personality traits are measured with different items on a 7-point Likert-scale.

[b)] Contacts to neighbors and childcare situation are measured on a 1-6 scale decreasing from good to bad and categorized into two groups.

[c)] Life satisfaction and expected ALMP probabilities are measured on a 0-10 scale increasing from low to high and categorized into three groups.

Table 3: Summary of Propensity Score Estimation: Sequential Inclusion of Usually Unobserved Variables

	Baseline specification: Socio-demographic characteristics, Household characteristics, Regional and seasonal information			*Baseline specification:* Socio-demographic characteristics, Household characteristics, Regional and seasonal information, Short-term labor market histories, Long-term labor market histories		
		$P-$value			$P-$value	
	Hitrate	F-test[1)]	SSVY-test[2)]	Hitrate	F-test[1)]	SSVY-test[2)]
Short-term training						
Baseline specification and ...	0.570		0.657	0.602		0.489
Personality traits	0.577	0.002	0.462	0.602	0.000	0.952
Intergenerational variables	0.569	0.439	0.354	0.602	0.512	0.894
Social network	0.572	0.033	0.570	0.603	0.045	0.480
Labor market flexibility	0.569	0.529	0.270	0.601	0.752	0.475
Life satisfaction	0.571	0.000	0.312	0.603	0.000	0.828
ALMP expectations	0.594	0.000	0.251	0.614	0.000	0.498
All of them	0.596	0.000	0.278	0.617	0.000	0.254
Long-term training						
Baseline specification and ...	0.574		0.010	0.626		0.406
Personality traits	0.585	0.002	0.657	0.635	0.005	0.096
Intergenerational variables	0.575	0.818	0.110	0.628	0.876	0.360
Social network	0.579	0.174	0.018	0.623	0.177	0.554
Labor market flexibility	0.578	0.270	0.019	0.628	0.374	0.417
Life satisfaction	0.584	0.002	0.569	0.626	0.001	0.679
ALMP expectations	0.660	0.000	0.064	0.674	0.000	0.595
All of them	0.662	0.000	0.797	0.675	0.000	0.509
Wage subsidies						
Baseline specification and ...	0.603		0.894	0.656		0.353
Personality traits	0.602	0.256	0.748	0.653	0.233	0.355
Intergenerational variables	0.608	0.383	0.203	0.660	0.294	0.620
Social networks	0.604	0.713	0.855	0.655	0.707	0.389
Labor market flexibility	0.607	0.067	0.670	0.655	0.111	0.586
Life satisfaction	0.607	0.005	0.398	0.655	0.006	0.434
ALMP expectations	0.614	0.000	0.537	0.662	0.000	0.475
All of them	0.623	0.000	0.306	0.669	0.000	0.215

Notes: Full estimation results are available in the appendix. $P-$values refer to 1) a F-test on joint significance when separately including each block of usually unobserved variables and 2) the specification test presented in Shaikh, Simonsen, Vytlacil, and Yildiz (2009) using a normal kernel with bandwidth $0.05n^{-1/8}$. Full estimation results for the main specifications can be found in Table A.2-A.4 in the Appendix.

Table 4: Consequences for Propensity Scores and Ranks

	Standard v. auxiliary	Extended v. standard
	(1)	(2)
Short-term training		
Propensity score correlation: Pearson's r	0.616	0.839
Rank correlation: Spearman's ρ	0.619	0.834
Absolute value of score difference: treated and matched controls$^{a)}$		
Mean	0.092	0.083
Median	0.078	0.071
Maximum	0.453	0.359
Distribution comparison: Wilcoxon signed-ranks test		
Participants	-1.832	9.356
	{0.067}	{0.000}
Non-participants	0.554	-6.968
	{0.580}	{0.000}
Long-term training		
Propensity score correlation: Pearson's r	0.463	0.740
Rank correlation: Spearman's ρ	0.503	0.749
Absolute value of score difference: treated and matched controls$^{a)}$		
Mean	0.084	0.099
Median	0.069	0.078
Maximum	0.541	0.418
Distribution comparison: Wilcoxon signed-ranks test		
Participants	-5.165	10.592
	{0.000}	{0.000}
Non-participants	4.061	-11.416
	{0.000}	{0.000}
Wage subsidies		
Propensity score correlation: Pearson's r	0.655	0.929
Rank correlation: Spearman's ρ	0.688	0.929
Absolute value of score difference: treated and matched controls$^{a)}$		
Mean	0.075	0.041
Median	0.055	0.031
Maximum	0.565	0.286
Distribution comparison: Wilcoxon signed-ranks test		
Participants	3.910	4.747
	{0.000}	{0.000}
Non-participants	-5.806	-6.269
	{0.000}	{0.000}

Notes: $P-$values are shown in brackets.
$^{a)}$This refers to: (1) the propensity score difference for the standard specification between participants and matched non-participants in the auxiliary specification, respectively (2) the propensity score difference for the extended specification between participants and matched non-participants in the standard specification.

Table 5: Consequences for Matching Quality: Mean Standardized Bias

	Unconditional raw difference	Propensity score specification		
		Standard (1)	Auxiliary (2)	Extended (3)
Short-term training				
Socio-demographic characteristics	5.10	2.33	1.98	2.29
Labor market histories	7.85	2.06	4.06	1.91
Personality traits	5.83	6.13	2.33	2.58
Intergenerational variables	5.42	2.25	1.33	2.12
Social network	6.70	4.59	1.79	3.39
Labor market flexibility	2.34	3.21	2.96	1.08
Life satisfaction	8.71	9.35	0.78	0.56
ALMP expectations	29.7	30.0	1.91	0.81
Total	6.59	3.32	2.57	2.11
Long-term training				
Socio-demographic characteristics	4.80	3.55	3.05	3.12
Labor market history	10.5	2.80	9.01	2.70
Personality traits	6.95	5.80	2.91	3.24
Intergenerational variables	7.95	1.66	4.94	5.26
Social network	6.50	5.98	4.97	1.57
Labor market flexibility	7.69	5.23	4.46	1.83
Life satisfaction	8.96	12.1	2.26	1.88
ALMP expectations	47.3	45.4	1.89	2.81
Total	7.96	4.59	4.94	3.02
Wage subsidies				
Socio-demographic characteristics	7.19	3.85	4.23	3.29
Labor market history	15.5	4.24	13.6	4.44
Personality traits	5.99	5.75	3.65	3.42
Intergenerational variables	8.56	3.20	2.38	1.77
Social network	0.62	1.99	1.17	1.06
Labor market flexibility	8.85	2.12	3.10	5.71
Life satisfaction	14.3	5.83	2.75	4.56
ALMP expectations	17.2	16.6	3.23	1.51
Total	9.95	4.29	6.71	3.54
Propensity score specification				
Personality traits			✓	✓
Inter-generational variables			✓	✓
Social network			✓	✓
Labor market flexibility			✓	✓
Life satisfaction			✓	✓
ALMP expectations			✓	✓
Socio-demographic characteristics		✓	✓	✓
Labor market history		✓		✓

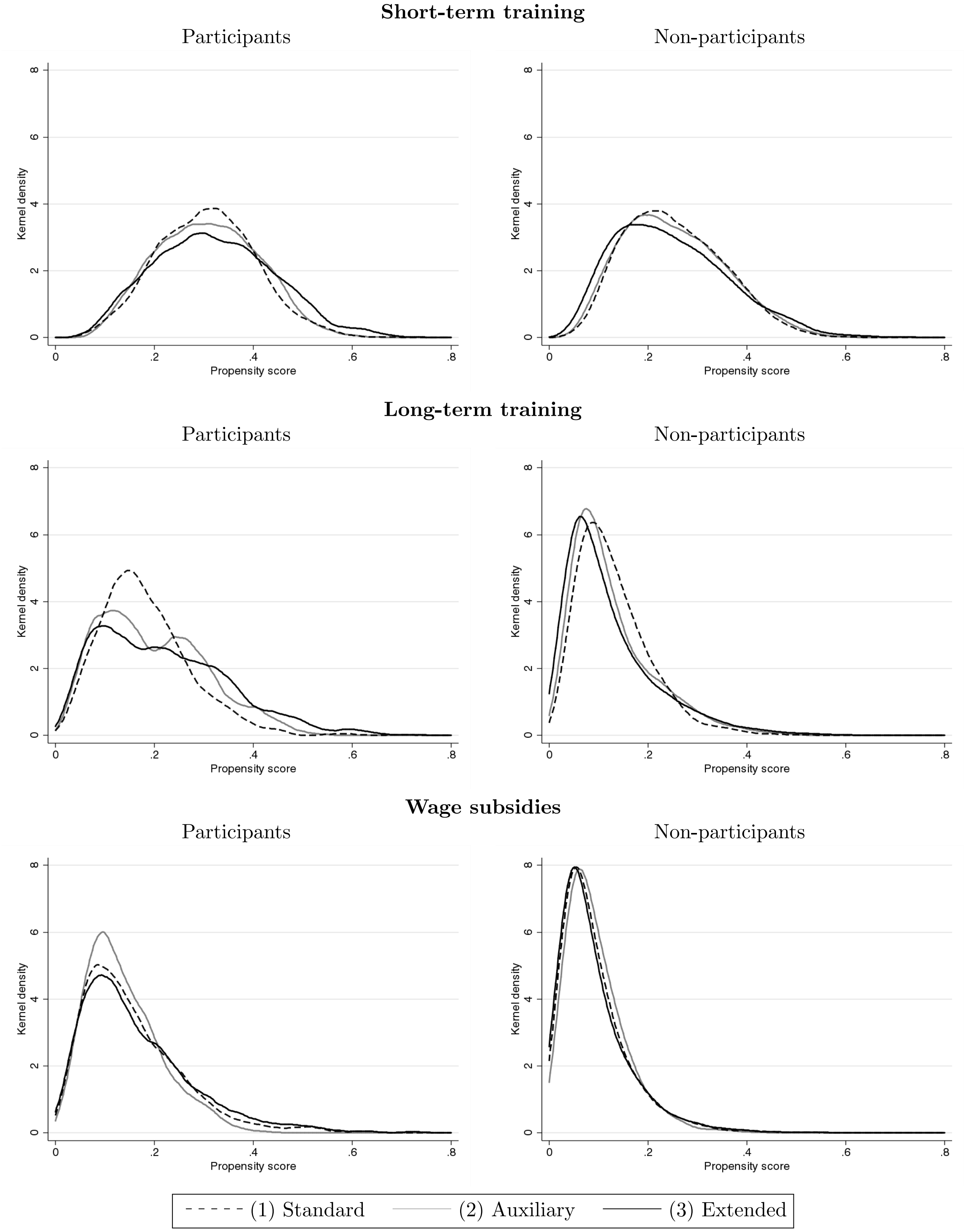
Notes: Reported are the mean standardized bias for each block of covariates calculated over the absolute standardized bias over all covariates in the block. The standardized bias is calculated as the difference of sample means for participants and non-participants as a percentage of the square root of the average of sample variances in both groups (Rosenbaum and Rubin, 1983).

Table 6: Matching Estimation Results: Consequences for the Average Treatment Effects on the Treated (ATT)

	Unconditional raw difference	Propensity score specification			Differences		
		Standard (1)	Auxiliary (2)	Extended (3)	Standard v. uncond.	Standard v. auxiliary	Extended v. standard
Short-term training							
Regular employed in t+30	0.008	-0.002	-0.004	-0.004	-0.0096***	0.0025	-0.0024
	(0.014)	(0.013)	(0.013)	(0.014)	(0.0036)	(0.0035)	(0.0028)
Cumulated earnings in € up to t+30	-2,416***	-1,905***	-2,129***	-1,886***	511***	224	19
	(647)	(585)	(489)	(535)	(174)	(157)	(133)
No. of observations	6,737	6,737	6,737	6,737			
Long-term training							
Regular employed in t+30	0.006	-0.009	-0.025	-0.017	-0.0143***	0.0164***	0.0089
	(0.019)	(0.021)	(0.021)	(0.022)	(0.0048)	(0.0062)	(0.0059)
Cumulated earnings in € up to t+30	-3,876***	-5,280***	-5,656***	-5,618***	-1,404***	-376	-338
	(938)	(823)	(846)	(859)	(331)	(327)	(323)
No. of observations	5,717	5,717	5,717	5,717			
Wage subsidies							
Regular employed in t+30	0.128***	0.095***	0.100***	0.087***	-0.0331***	-0.0051	-0.0075*
	(0.022)	(0.022)	(0.022)	(0.023)	(0.0063)	(0.0058)	(0.0045)
Cumulated earnings in € up to t+30	4,783***	3,924***	3,723***	3,691***	-859***	201	233
	(1,105)	(805)	(830)	(839)	(284)	(262)	(194)
No. of observations	5,476	5,476	5,476	5,476			
Propensity score specification							
Personality traits			✓	✓			
Inter-generational variables			✓	✓			
Social network			✓	✓			
Labor market flexibility			✓	✓			
Life satisfaction			✓	✓			
ALMP expectations			✓	✓			
Socio-demographic characteristics		✓	✓	✓			
Labor market history		✓		✓			

Note: Depicted are estimated average treatment effects on the treated as the difference in mean outcomes between participants and matched non-participants using epanechnikov kernel propensity score matching with bandwidth 0.06. Standard errors are in parentheses and based on bootstrapping with 999 replications. Standard errors for the differences in ATT's are based on bootstrapped robust Hausman tests with 999 replications (see Cameron and Trivedi, 2010, for details). ***/**/* indicate statistically significance at the 1/5/10%-level.

Figure 1: Propensity Score Distribution



Note: Depicted are epanechnikov kernel densities (bandwith=0.06) of the propensity score after matching on the four propensity score specifications.

Supplementary Appendix

Unobservable, but Unimportant? The Relevance of Usually Unobserved Variables for the Evaluation of Labor Market Policies

Marco Caliendo* **Robert Mahlstedt†**
Oscar A. Mitnik‡

February 18, 2017

The supplementary appendix provides additional information and is not intended to be published but will be made available online.

*University of Potsdam, IZA Bonn, DIW Berlin, IAB Nuremberg; e-mail: caliendo@empwifo.uni-potsdam.de. Corresponding address: University of Potsdam, Chair of Empirical Economics, August-Bebel-Str. 89, 14482 Potsdam, Germany. Tel: +49 331 977 3225. Fax: +49 331 977 3210.
†IZA Bonn, University of Potsdam; e-mail: mahlstedt@iza.org.
‡Inter-American Development Bank and IZA Bonn; e-mail: oscar@mitnik.net.

Table A.1: Overview - Control Variables and Propensity Score Specifications

	Propensity score specification		
	Standard (1)	Auxiliary (2)	Extended (3)
1) Usually unobserved variables			
A. Personality traits Openness, Conscientiousness, Extraversion, Neuroticism, Locus of control		✓	✓
B. Intergenerational variables Father has A-level qualification (upper sec. degree), Father's current age, Father was employed when person aged 15, Father was blue-collar worker when person aged 15		✓	✓
C. Social Network Number of good friends outside the family, Contact to neighbors		✓	✓
D. Labor market flexibility Childcare situation, Car-ownership		✓	✓
E. Life satisfaction Life satisfaction: low, medium, high		✓	✓
F. ALMP expectations Expected ALMP participation probability: low, medium, high		✓	✓
2) Socio-demographic/ baseline variables			
A. Individual and household characteristics Gender, Age, Migration background, School leaving degree, Level of higher education, Marital status, German citizenship, Number of children, Health problems, Searching for full- or part-time employment, Employment status of partner	✓	✓	✓
B. Regional and seasonal information Living in West-Germany, Local unemployment rate, Month of entry into unemployment, Time between entry into UE and interview	✓	✓	✓
3) Labor market history			
A. Short-term labor market history Employment status before entry into unemployment, Last daily wage, Last job was full-time employment, Laid off by last employer, Time with last employer, Duration of last unemployment spell Months in employment/ unemployment/ ALMP program/ out of labor force in last 6 months/ 24 months, Number of employers/ program participation/ unemployment spells/ out of labor force spells in last 24 months	✓		✓
B. Long-term labor market history Months in employment/ unemployment/ ALMP program/ out of labor force in last 10 years, Number of employers/ program participation/ unemployment spells/ out of labor force spells in last 10 years	✓		✓

Table A.2: Propensity Score Estimation: Short-term Training

	Standard (1)		Auxiliary (2)		Extended (3)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Personality traits[a)]						
Openness (standardized)			-0.007	(0.006)	-0.005	(0.006)
Conscientiousness (standardized)			0.005	(0.006)	0.003	(0.006)
Extraversion (standardized)			-0.007	(0.006)	-0.008	(0.006)
Neuroticism (standardized)			-0.007	(0.006)	-0.008	(0.006)
Locus of control (standardized)			-0.017***	(0.006)	-0.020***	(0.006)
Intergenerational variables						
Father has upper sec. school leaving degree			-0.023	(0.017)	-0.020	(0.017)
Father was employed when person aged 15			-0.001	(0.016)	-0.000	(0.016)
Father's current age (Ref.: already passed away)			ref.		ref.	
$\leq$ 60 years			-0.017	(0.017)	-0.017	(0.017)
$>$ 60 years			-0.001	(0.014)	-0.005	(0.014)
Father was blue-collar worker when person aged 15			-0.001	(0.013)	0.002	(0.013)
Social network[b)]						
No. of good friends outside family: Less than two			0.045**	(0.019)	0.041**	(0.018)
Contacts to neighbors: Good (1-3)			-0.002	(0.012)	-0.000	(0.012)
Labor market flexibility						
Childcare situation: Bad (4-6)			0.010	(0.025)	0.000	(0.025)
Car-ownership			0.020*	(0.012)	0.017	(0.012)
Life satisfaction[c)]						
Life satisfaction (Ref.: Medium (4-6))			ref.		ref.	
Low (0-3)			-0.003	(0.019)	-0.006	(0.019)
High (7-10)			-0.035***	(0.011)	-0.037***	(0.011)
ALMP expectations[c)]						
Expected ALMP probability (Ref.: Medium (4-6))			ref.		ref.	
Low (0-3)			-0.085***	(0.012)	-0.078***	(0.012)
High (7-10)			0.054***	(0.015)	0.053***	(0.015)
Socio-demographic characteristics						
Female	0.037***	(0.013)	0.045***	(0.013)	0.033**	(0.013)
Age (Ref.: 16-24 years)	ref.		ref.		ref.	
25-34 years	-0.013	(0.019)	-0.022	(0.017)	-0.020	(0.019)
35-44 years	-0.024	(0.021)	-0.045**	(0.020)	-0.045**	(0.023)
45-55 years	-0.012	(0.023)	-0.039*	(0.022)	-0.040	(0.025)
Married or cohabiting	0.019	(0.016)	0.023	(0.016)	0.018	(0.016)
German citizenship	-0.013	(0.029)	-0.008	(0.029)	-0.009	(0.029)
Migration background	0.003	(0.020)	-0.007	(0.019)	-0.004	(0.019)
Children (Ref.: None)	ref.		ref.		ref.	
One child	-0.005	(0.015)	-0.000	(0.015)	-0.002	(0.015)
Two children or more	0.001	(0.019)	0.001	(0.019)	-0.001	(0.019)
School leaving degree (Ref.: None)	ref.		ref.		ref.	
Lower sec. degree	-0.018	(0.035)	-0.009	(0.035)	-0.008	(0.034)
Middle sec. degree	-0.035	(0.034)	-0.019	(0.035)	-0.020	(0.034)
Upper sec. degree	-0.098***	(0.030)	-0.079**	(0.032)	-0.073**	(0.032)
Higher education (Ref.: None)	ref.		ref.		ref.	
Int. or ext. vocational training	0.023	(0.019)	0.036*	(0.019)	0.028	(0.019)
University degree	-0.029	(0.022)	-0.017	(0.022)	-0.015	(0.023)
Month of entry into unemployment (Ref.: June)	ref.		ref.		ref.	
July	-0.015	(0.030)	-0.018	(0.030)	-0.022	(0.030)
August	-0.029	(0.030)	-0.036	(0.028)	-0.042	(0.029)
September	-0.026	(0.030)	-0.015	(0.030)	-0.023	(0.031)
October	0.007	(0.029)	0.006	(0.028)	0.001	(0.029)
November	-0.034	(0.027)	-0.038	(0.026)	-0.038	(0.027)
December	0.005	(0.032)	-0.004	(0.031)	0.001	(0.032)
January	-0.038	(0.031)	-0.042	(0.030)	-0.048	(0.031)
February	0.024	(0.033)	0.031	(0.033)	0.016	(0.033)
March	-0.008	(0.032)	0.010	(0.033)	-0.010	(0.032)
April	-0.074***	(0.026)	-0.057**	(0.027)	-0.074***	(0.026)
May	-0.039	(0.028)	-0.023	(0.029)	-0.044	(0.028)

Continued on next page.

Continued from previous page.

Time between entry into UE and interview (Ref.: 7 weeks)	ref.		ref.		ref.	
8 weeks	0.009	(0.040)	-0.004	(0.039)	-0.006	(0.040)
9 weeks	0.040	(0.044)	0.024	(0.043)	0.023	(0.043)
10 weeks	0.028	(0.044)	0.013	(0.043)	0.014	(0.044)
11 weeks	0.030	(0.046)	0.022	(0.046)	0.018	(0.046)
12 weeks	0.041	(0.051)	0.026	(0.050)	0.027	(0.050)
13 weeks	0.006	(0.052)	-0.012	(0.051)	-0.016	(0.051)
14 weeks	0.036	(0.051)	0.028	(0.051)	0.019	(0.050)
Health restriction or disability	-0.063***	(0.020)	-0.069***	(0.019)	-0.068***	(0.019)
Job search (Ref.: Full- or part-time employment)	ref.		ref.		ref.	
Full-time employment only	0.088***	(0.013)	0.086***	(0.013)	0.082***	(0.013)
Part-time employment only	0.044**	(0.020)	0.064***	(0.020)	0.050**	(0.021)
Employment status partner (Ref.: No partner)	ref.		ref.		ref.	
Full-time employed	-0.017	(0.015)	-0.006	(0.015)	-0.005	(0.015)
Part-time employed	-0.011	(0.025)	-0.008	(0.024)	-0.005	(0.025)
Education	-0.021	(0.023)	-0.016	(0.022)	-0.012	(0.023)
Unemployment	-0.050*	(0.028)	-0.043	(0.027)	-0.041	(0.027)
Other	0.002	(0.025)	0.015	(0.025)	0.011	(0.025)
Region (Ref.: West-Germany: UE rate 0-3%)	ref.		ref.		ref.	
West-Germany: UE rate 4-6%	0.085	(0.083)	0.063	(0.084)	0.062	(0.085)
West-Germany: UE rate 7-9%	0.072	(0.081)	0.063	(0.084)	0.056	(0.084)
West-Germany: UE rate $\geq$ 10%	0.054	(0.078)	0.045	(0.082)	0.038	(0.081)
East-Germany: UE rate 9-12%	0.054	(0.080)	0.043	(0.083)	0.041	(0.084)
East-Germany: UE rate 13-14%	0.065	(0.081)	0.061	(0.085)	0.056	(0.085)
East-Germany: UE rate 15-16%	0.014	(0.071)	0.004	(0.074)	-0.001	(0.074)
East-Germany: UE rate $\geq$ 17%	0.072	(0.083)	0.066	(0.087)	0.059	(0.086)
Short-term labor market history						
Employment status before UE (Ref.: Regular employed)	ref.				ref.	
Subsidized employed	0.014	(0.022)			0.017	(0.022)
School, apprentice, military, etc.	-0.007	(0.019)			-0.004	(0.019)
Parental leave	0.086**	(0.036)			0.084**	(0.035)
Months employed in last 6 months	-0.004	(0.006)			-0.004	(0.006)
Months unemployed in last 6 months	0.000	(0.007)			0.001	(0.007)
Month out of labor force in last 6 months	-0.002	(0.007)			-0.002	(0.007)
Months employed in last 24 months	-0.001	(0.002)			-0.002	(0.002)
Months unemployed in last 24 months	-0.011***	(0.002)			-0.011***	(0.002)
Months out of labor force in last 24 months	-0.003	(0.002)			-0.003	(0.002)
No. of employers in last 24 months	-0.011	(0.007)			-0.011	(0.007)
No. of unemployment spells in last 24 months	0.016*	(0.009)			0.014	(0.009)
No. of ALMP programs in last 24 months	0.035***	(0.011)			0.033***	(0.011)
No. of out of labor force spells in last 24 months	-0.023*	(0.012)			-0.024*	(0.012)
Last daily income in €	-0.001***	(0.000)			-0.000**	(0.000)
Last job: Full-time employment	-0.057**	(0.025)			-0.058**	(0.024)
Last job: Laid off by employer	0.003	(0.018)			0.003	(0.018)
Long-term labor market history						
Months employed in last 10 years	0.002***	(0.001)			0.002***	(0.001)
Months unemployed in last 10 years	0.000***	(0.000)			0.000***	(0.000)
Months out of labor force in last 10 years	0.002***	(0.001)			0.002***	(0.001)
No. of employers in last 10 years	0.003	(0.003)			0.003	(0.003)
No. of unemployment spells in last 10 years	-0.015***	(0.004)			-0.015***	(0.004)
No. of ALMP programs in last 10 years	0.017**	(0.007)			0.015**	(0.007)
No. of out of labor force spells in last 10 years	0.000	(0.005)			0.002	(0.005)
Time with last employer in 100 days	0.000	(0.001)			0.000	(0.001)
Duration of last unemployment spell in 100 days	0.017	(0.067)			0.026	(0.066)
Months in ALMP programs in last 10 years	0.000	(0.001)			0.000	(0.001)
Observations	6,737		6,737		6,737	
log-Likelihood	-3729.545		-3716.352		-3653.668	
Mean value	0.268		0.268		0.268	
Pseudo-R^2	0.047		0.050		0.066	

Note: Depicted are average marginal effects based on logit models estimating the participation probability. ***/**/* indicate statistically significance at the 1%/5%/10%-level. Standard errors in parenthesis.
[a] Personality traits are measured with different items on a 7-Point Likert-scale.
[b] Contacts to neighbors and the childcare situation are measured on a scale from 1 (very good) to 6 (very bad).
[c] Life satisfaction and expected ALMP probabilities are measured on a scale from 0 (very low) to 10 (very high).

Table A.3: Propensity Score Estimation: Long-term Training

	Standard (1)		Auxiliary (2)		Extended (3)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Personality traits$^{a)}$						
Openness (standardized)			-0.009*	(0.005)	-0.008	(0.005)
Conscientiousness (standardized)			0.009*	(0.005)	0.008	(0.005)
Extraversion (standardized)			-0.017***	(0.005)	-0.016***	(0.005)
Neuroticism (standardized)			-0.005	(0.005)	-0.005	(0.005)
Locus of control (standardized)			-0.003	(0.005)	-0.004	(0.005)
Intergenerational variables						
Father has upper sec. school leaving degree			-0.007	(0.014)	-0.005	(0.014)
Father was employed when person aged 15			-0.002	(0.014)	0.001	(0.014)
Father's current age (Ref.: already passed away)			ref.		ref.	
$\leq$ 60 years			0.002	(0.015)	0.002	(0.015)
$>$ 60 years			0.015	(0.012)	0.013	(0.012)
Father was blue-collar worker when person aged 15			-0.006	(0.011)	-0.007	(0.011)
Social Network$^{b)}$						
No. of good friends outside family: Less than two			0.010	(0.016)	0.008	(0.016)
Contacts to neighbors: Good (1-3)			-0.014	(0.010)	-0.014	(0.010)
Labor Market Flexibility						
Childcare situation: Bad (4-6)			0.005	(0.020)	-0.010	(0.019)
Car-ownership			0.020*	(0.010)	0.018*	(0.010)
Life satisfaction$^{c)}$						
Life satisfaction (Ref.: Medium (4-6))			ref.		ref.	
Low (0-3)			-0.006	(0.017)	-0.010	(0.017)
High (7-10)			-0.030***	(0.009)	-0.032***	(0.009)
ALMP expectations$^{c)}$						
Expected ALMP probability (Ref.: Medium (4-6))			ref.		ref.	
Low (0-3)			-0.039***	(0.009)	-0.038***	(0.009)
High (7-10)			0.106***	(0.017)	0.102***	(0.017)
Socio-demographic characteristics						
Female	0.021*	(0.011)	0.030***	(0.011)	0.022*	(0.012)
Age (Ref.: 16-24 years)	ref.		ref.		ref.	
25-34 years	0.047***	(0.017)	0.039**	(0.016)	0.043**	(0.018)
35-44 years	0.079***	(0.022)	0.052***	(0.020)	0.063***	(0.024)
45-55 years	0.090***	(0.025)	0.061***	(0.022)	0.074***	(0.027)
Married or cohabiting	-0.015	(0.013)	-0.002	(0.013)	-0.012	(0.013)
German citizenship	0.031	(0.025)	0.030	(0.025)	0.029	(0.025)
Migration background	0.040**	(0.018)	0.026	(0.017)	0.030*	(0.018)
Children (Ref.: None)	ref.		ref.		ref.	
One child	0.006	(0.013)	0.008	(0.013)	0.005	(0.014)
Two children or more	0.028	(0.017)	0.035**	(0.017)	0.024	(0.018)
School leaving degree (Ref.: None)	ref.		ref.		ref.	
Lower sec. degree	0.032	(0.035)	0.044	(0.036)	0.042	(0.036)
Middle sec. degree	0.041	(0.037)	0.053	(0.038)	0.048	(0.037)
Upper sec. degree	0.034	(0.037)	0.049	(0.038)	0.045	(0.038)
Higher education (Ref.: None)	ref.		ref.		ref.	
Int. or ext. vocational training	-0.012	(0.016)	-0.008	(0.016)	-0.014	(0.016)
University degree	-0.014	(0.020)	-0.009	(0.019)	-0.014	(0.020)
Month of entry into unemployment (Ref.: June)	ref.		ref.		ref.	
July	-0.038*	(0.022)	-0.038*	(0.022)	-0.035	(0.022)
August	-0.016	(0.025)	-0.028	(0.023)	-0.020	(0.024)
September	0.018	(0.029)	0.023	(0.028)	0.024	(0.029)
October	-0.001	(0.025)	-0.004	(0.024)	-0.000	(0.025)
November	0.015	(0.026)	0.003	(0.024)	0.019	(0.026)
December	0.005	(0.028)	-0.005	(0.026)	0.013	(0.028)
January	-0.032	(0.024)	-0.039*	(0.022)	-0.032	(0.023)
February	0.009	(0.028)	0.010	(0.028)	0.008	(0.028)
March	-0.018	(0.025)	-0.006	(0.026)	-0.010	(0.026)
April	-0.003	(0.025)	0.004	(0.026)	0.001	(0.025)
May	0.026	(0.028)	0.034	(0.028)	0.027	(0.027)

Continued on next page.

Continued from previous page.

Time between entry into UE and interview (Ref.: 7 weeks)	ref.		ref.		ref.	
8 weeks	0.000	(0.036)	-0.008	(0.035)	-0.014	(0.034)
9 weeks	0.011	(0.038)	-0.005	(0.037)	-0.009	(0.036)
10 weeks	0.022	(0.042)	0.014	(0.041)	0.011	(0.040)
11 weeks	0.003	(0.039)	-0.002	(0.040)	-0.009	(0.038)
12 weeks	-0.014	(0.038)	-0.020	(0.039)	-0.025	(0.038)
13 weeks	-0.017	(0.041)	-0.025	(0.041)	-0.036	(0.039)
14 weeks	-0.014	(0.040)	-0.017	(0.041)	-0.024	(0.040)
Health restriction or disability	-0.030*	(0.016)	-0.038**	(0.016)	-0.030*	(0.017)
Job search (Ref.: Full- or part-time employment)	ref.		ref.		ref.	
Full-time employment only	0.055***	(0.013)	0.052***	(0.012)	0.047***	(0.012)
Part-time employment only	0.006	(0.015)	0.020	(0.016)	0.010	(0.016)
Employment status partner (Ref.: No partner)	ref.		ref.		ref.	
Full-time employed	-0.006	(0.013)	-0.002	(0.013)	-0.000	(0.013)
Part-time employed	-0.042**	(0.018)	-0.036**	(0.018)	-0.035*	(0.019)
Education	-0.011	(0.021)	-0.014	(0.021)	-0.007	(0.021)
Unemployment	0.009	(0.027)	0.006	(0.026)	0.010	(0.026)
Other	-0.010	(0.022)	-0.004	(0.022)	-0.004	(0.022)
Region (Ref.: West-Germany: UE rate 0-3%)	ref.		ref.		ref.	
West-Germany: UE rate 4-6%	-0.060	(0.054)	-0.066	(0.049)	-0.068	(0.048)
West-Germany: UE rate 7-9%	-0.079	(0.049)	-0.075	(0.046)	-0.082*	(0.045)
West-Germany: UE rate $\geq$ 10%	-0.062	(0.054)	-0.055	(0.052)	-0.063	(0.050)
East-Germany: UE rate 9-12%	-0.113***	(0.042)	-0.114***	(0.038)	-0.115***	(0.040)
East-Germany: UE rate 13-14%	-0.084*	(0.050)	-0.080*	(0.047)	-0.078	(0.048)
East-Germany: UE rate 15-16%	-0.078	(0.051)	-0.077	(0.048)	-0.076	(0.049)
East-Germany: UE rate $\geq$ 17%	-0.055	(0.058)	-0.048	(0.056)	-0.050	(0.055)
Short-term labor market history						
Employment status before UE (Ref.: Regular employed)	ref.				ref.	
Subsidized employed	-0.030*	(0.017)			-0.029*	(0.017)
School, apprentice, military, etc.	-0.007	(0.018)			-0.008	(0.018)
Parental leave	0.106***	(0.035)			0.108***	(0.035)
Months employed in last 6 months	-0.007	(0.005)			-0.007	(0.005)
Months unemployed in last 6 months	0.005	(0.006)			0.004	(0.006)
Month out of labor force in last 6 months	-0.006	(0.006)			-0.008	(0.006)
Months employed in last 24 months	0.001	(0.001)			0.000	(0.002)
Months unemployed in last 24 months	-0.006***	(0.002)			-0.005***	(0.002)
Months out of labor force in last 24 months	0.000	(0.002)			0.000	(0.002)
No. of employers in last 24 months	-0.002	(0.006)			-0.001	(0.006)
No. of unemployment spells in last 24 months	0.007	(0.008)			0.005	(0.008)
No. of ALMP programs in last 24 months	0.035***	(0.009)			0.034***	(0.009)
No. of out of labor force spells in last 24 months	-0.025**	(0.011)			-0.029***	(0.011)
Last daily income in €	-0.000	(0.000)			0.000	(0.000)
Last job: Full-time employment	0.005	(0.023)			0.002	(0.023)
Last job: Laid off by employer	0.013	(0.015)			0.012	(0.015)
Long-term labor market history						
Months employed in last 10 years	-0.000	(0.000)			0.000	(0.000)
Months unemployed in last 10 years	-0.000*	(0.000)			-0.000	(0.000)
Months out of labor force in last 10 years	-0.000	(0.000)			-0.000	(0.000)
No. of employers in last 10 years	0.001	(0.003)			0.000	(0.003)
No. of unemployment spells in last 10 years	-0.008**	(0.004)			-0.009**	(0.004)
No. of ALMP programs in last 10 years	0.005	(0.006)			0.004	(0.006)
No. of out of labor force spells in last 10 years	-0.003	(0.005)			-0.001	(0.005)
Time with last employer in 100 days	0.001	(0.001)			0.001	(0.001)
Duration of last unemployment spell in 100 days	0.071	(0.053)			0.085	(0.053)
Months in ALMP programs in last 10 years	0.001	(0.001)			0.001	(0.001)
Observations	5,717		5,717		5,717	
log-Likelihood	-2144.481		-2099.245		-2040.801	
Mean value	0.137		0.137		0.137	
Pseudo-R^2	0.061		0.081		0.106	

Note: Depicted are average marginal effects based on logit models estimating the participation probability. ***/**/* indicate statistically significance at the 1%/5%/10%-level. Standard errors in parenthesis.

[a)] Personality traits are measured with different items on a 7-Point Likert-scale.

[b)] Contacts to neighbors and the childcare situation are measured on a scale from 1 (very good) to 6 (very bad).

[c)] Life satisfaction and expected ALMP probabilities are measured on a scale from 0 (very low) to 10 (very high).

Table A.4: Propensity Score Estimation: Wage Subsidies

	Standard (1)		Auxiliary (2)		Extended (3)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Personality traits$^{a)}$						
Openness (standardized)			0.003	(0.005)	0.006	(0.005)
Conscientiousness (standardized)			0.008	(0.005)	0.007	(0.005)
Extraversion (standardized)			-0.005	(0.005)	-0.005	(0.005)
Neuroticism (standardized)			-0.007	(0.004)	-0.007	(0.004)
Locus of control (standardized)			-0.003	(0.005)	-0.003	(0.005)
Intergenerational variables						
Father has upper sec. school leaving degree			-0.014	(0.013)	-0.012	(0.013)
Father was employed when person aged 15			0.012	(0.013)	0.014	(0.013)
Father's current age (Ref.: already passed away)			ref.		ref.	
$\leq$ 60 years			-0.004	(0.014)	-0.003	(0.015)
$>$ 60 years			-0.016	(0.010)	-0.018*	(0.010)
Father was blue-collar worker when person aged 15			0.000	(0.010)	0.000	(0.010)
Social network$^{b)}$						
No. of good friends outside family: Less than two			-0.010	(0.013)	-0.009	(0.013)
Contacts to neighbors: Good (1-3)			0.000	(0.010)	0.002	(0.010)
Labor market flexibility						
Childcare situation: Bad (4-6)			-0.012	(0.018)	-0.018	(0.017)
Car-ownership			0.022**	(0.010)	0.018*	(0.010)
Life satisfaction$^{c)}$						
Life satisfaction (Ref.: Medium (4-6))			ref.		ref.	
Low (0-3)			0.011	(0.016)	0.008	(0.016)
High (7-10)			-0.026***	(0.008)	-0.026***	(0.008)
ALMP expectations$^{c)}$						
Expected ALMP probability (Ref.: Medium (4-6))			ref.		ref.	
Low (0-3)			-0.024***	(0.009)	-0.019**	(0.009)
High (7-10)			0.020	(0.013)	0.021*	(0.013)
Socio-demographic characteristics						
Female	0.003	(0.009)	0.013	(0.010)	0.005	(0.010)
Age (Ref.: 16-24 years)	ref.		ref.		ref.	
25-34 years	0.013	(0.014)	0.007	(0.012)	0.013	(0.014)
35-44 years	0.031*	(0.018)	0.019	(0.017)	0.032	(0.021)
45-55 years	0.086***	(0.025)	0.067***	(0.024)	0.077***	(0.029)
Married or cohabiting	-0.012	(0.011)	-0.009	(0.011)	-0.011	(0.011)
German citizenship	0.003	(0.024)	0.003	(0.024)	0.001	(0.024)
Migration background	0.012	(0.017)	0.012	(0.017)	0.015	(0.017)
Children (Ref.: None)	ref.		ref.		ref.	
One child	0.004	(0.011)	0.006	(0.012)	0.006	(0.012)
Two children or more	0.028*	(0.016)	0.032*	(0.017)	0.030*	(0.017)
School leaving degree (Ref.: None)	ref.		ref.		ref.	
Lower sec. degree	-0.039	(0.026)	-0.039	(0.026)	-0.038	(0.026)
Middle sec. degree	-0.047*	(0.024)	-0.046*	(0.024)	-0.046*	(0.025)
Upper sec. degree	-0.050**	(0.025)	-0.054**	(0.024)	-0.046*	(0.026)
Higher education (Ref.: None)	ref.		ref.		ref.	
Int. or ext. vocational training	0.021	(0.017)	0.026	(0.017)	0.020	(0.017)
University degree	0.010	(0.019)	0.007	(0.018)	0.011	(0.019)
Month of entry into unemployment (Ref.: June)	ref.		ref.		ref.	
July	-0.034	(0.022)	-0.027	(0.022)	-0.038*	(0.022)
August	-0.008	(0.026)	-0.000	(0.026)	-0.010	(0.026)
September	-0.014	(0.025)	-0.005	(0.025)	-0.012	(0.026)
October	-0.009	(0.024)	0.001	(0.024)	-0.007	(0.025)
November	-0.035*	(0.019)	-0.033*	(0.018)	-0.037*	(0.020)
December	-0.033	(0.023)	-0.035*	(0.020)	-0.034	(0.023)
January	-0.062***	(0.018)	-0.060***	(0.017)	-0.065***	(0.018)
February	-0.013	(0.025)	-0.003	(0.026)	-0.015	(0.026)
March	-0.015	(0.025)	0.002	(0.027)	-0.014	(0.026)
April	-0.045**	(0.018)	-0.038**	(0.019)	-0.049***	(0.018)
May	-0.014	(0.024)	0.000	(0.025)	-0.015	(0.024)

Continued on next page.

Continued from previous page.

Time between entry into UE and interview (Ref.: 7 weeks)	ref.		ref.		ref.	
8 weeks	-0.022	(0.026)	-0.034	(0.027)	-0.029	(0.026)
9 weeks	0.004	(0.033)	-0.010	(0.033)	-0.003	(0.032)
10 weeks	-0.001	(0.033)	-0.019	(0.032)	-0.008	(0.033)
11 weeks	-0.008	(0.032)	-0.019	(0.034)	-0.016	(0.032)
12 weeks	0.026	(0.043)	0.013	(0.045)	0.019	(0.043)
13 weeks	-0.029	(0.032)	-0.040	(0.033)	-0.040	(0.031)
14 weeks	-0.019	(0.032)	-0.029	(0.034)	-0.027	(0.032)
Health restriction or disability	-0.048***	(0.012)	-0.053***	(0.011)	-0.048***	(0.013)
Job search (Ref.: Full- or part-time employment)	ref.		ref.		ref.	
Full-time employment only	0.060***	(0.012)	0.060***	(0.012)	0.056***	(0.012)
Part-time employment only	0.012	(0.014)	0.021	(0.014)	0.017	(0.015)
Employment status partner (Ref.: No partner)	ref.		ref.		ref.	
Full-time employed	0.012	(0.012)	0.013	(0.012)	0.015	(0.012)
Part-time employed	0.006	(0.018)	0.009	(0.018)	0.007	(0.018)
Education	0.015	(0.020)	0.014	(0.020)	0.019	(0.021)
Unemployment	0.002	(0.023)	-0.002	(0.022)	0.003	(0.023)
Other	-0.022	(0.015)	-0.021	(0.015)	-0.021	(0.016)
Region (Ref.: West-Germany: UE rate 0-3%)	ref.		ref.		ref.	
West-Germany: UE rate 4-6%	-0.086**	(0.036)	-0.078**	(0.035)	-0.090**	(0.036)
West-Germany: UE rate 7-9%	-0.086**	(0.036)	-0.072*	(0.037)	-0.089**	(0.037)
West-Germany: UE rate $\geq$ 10%	-0.081**	(0.038)	-0.068*	(0.039)	-0.084**	(0.039)
East-Germany: UE rate 9-12%	-0.071*	(0.043)	-0.059	(0.044)	-0.072	(0.045)
East-Germany: UE rate 13-14%	-0.058	(0.046)	-0.044	(0.047)	-0.061	(0.048)
East-Germany: UE rate 15-16%	-0.066	(0.043)	-0.048	(0.046)	-0.066	(0.045)
East-Germany: UE rate $\geq$ 17%	-0.049	(0.049)	-0.033	(0.051)	-0.046	(0.052)
Short-term labor market history						
Employment status before UE (Ref.: Regular employed)	ref.				ref.	
Subsidized employed	0.019	(0.017)			0.020	(0.017)
School, apprentice, military, etc.	0.007	(0.017)			0.008	(0.017)
Parental leave	0.033	(0.033)			0.038	(0.034)
Months employed in last 6 months	-0.000	(0.005)			-0.000	(0.005)
Months unemployed in last 6 months	0.004	(0.005)			0.004	(0.005)
Month out of labor force in last 6 months	-0.003	(0.006)			-0.004	(0.006)
Months employed in last 24 months	-0.001	(0.001)			-0.001	(0.001)
Months unemployed in last 24 months	-0.005***	(0.002)			-0.005***	(0.002)
Months out of labor force in last 24 months	-0.001	(0.002)			-0.001	(0.002)
No. of employers in last 24 months	-0.005	(0.005)			-0.006	(0.005)
No. of unemployment spells in last 24 months	0.015**	(0.007)			0.015**	(0.007)
No. of ALMP programs in last 24 months	0.018**	(0.008)			0.016**	(0.008)
No. of out of labor force spells in last 24 months	-0.020*	(0.010)			-0.022**	(0.011)
Last daily income in €	-0.001***	(0.000)			-0.001***	(0.000)
Last job: Full-time employment	-0.003	(0.020)			-0.002	(0.020)
Last job: Laid off by employer	-0.011	(0.013)			-0.009	(0.013)
Long-term labor market history						
Months employed in last 10 years	0.001**	(0.000)			0.001**	(0.000)
Months unemployed in last 10 years	-0.000	(0.000)			-0.000	(0.000)
Months out of labor force in last 10 years	0.001	(0.000)			0.001	(0.000)
No. of employers in last 10 years	0.007***	(0.002)			0.007***	(0.002)
No. of unemployment spells in last 10 years	-0.009***	(0.003)			-0.010***	(0.003)
No. of ALMP programs in last 10 years	0.012**	(0.005)			0.012**	(0.005)
No. of out of labor force spells in last 10 years	-0.004	(0.004)			-0.003	(0.004)
Time with last employer in 100 days	-0.001	(0.001)			-0.002	(0.001)
Duration of last unemployment spell in 100 days	0.027	(0.045)			0.035	(0.045)
Months in ALMP programs in last 10 years	0.000	(0.001)			0.000	(0.001)
Observations	5,476		5,476		5,476	
log-Likelihood	-1614.328		-1654.380		-1592.547	
Mean value	0.099		0.099		0.099	
Pseudo-R^2	0.087		0.064		0.099	

Note: Depicted are average marginal effects based on logit models estimating the participation probability. ***/**/* indicate statistically significance at the 1%/5%/10%-level. Standard errors in parenthesis.
[a)] Personality traits are measured with different items on a 7-Point Likert-scale.
[b)] Contacts to neighbors and the childcare situation are measured on a scale from 1 (very good) to 6 (very bad).
[c)] Life satisfaction and expected ALMP probabilities are measured on a scale from 0 (very low) to 10 (very high).

Table A.5: Consequences for the Average Treatment Effects on the Treated (ATT) for Different Matching/Weighting Estimators - Short-term Training

	Unconditional raw difference	Propensity score specification		
		Standard (1)	Auxiliary (2)	Extended (3)
Regular employed 30 months after entry (t+30)				
Regression	0.008	-0.007	-0.001	-0.005
	(0.014)	(0.013)	(0.013)	(0.013)
Nearest Neighbor (1:1)		-0.031	-0.007	0.019
		(0.022)	(0.022)	(0.022)
Radius 1 (caliper=0.1)		-0.0002	-0.002	-0.003
		(0.013)	(0.013)	(0.013)
Radius 2 (regression adjustment)		-0.026	-0.009	0.006
		(0.020)	(0.020)	(0.020)
Inverse Probability Weighting		-0.004	0.0005	-0.003
		(0.013)	(0.013)	(0.014)
Kernel 1 (bw=0.02)		-0.004	-0.002	-0.003
		(0.013)	(0.013)	(0.014)
Kernel 2 (bw=0.06)		-0.004	-0.002	-0.004
		(0.013)	(0.013)	(0.014)
Kernel 3 (bw=0.2)		-0.0001	0.002	-0.001
		(0.013)	(0.013)	(0.013)
Cumulated earnings in € up to t+30				
Regression	-2,416***	-2,231***	-1,831***	-1,945***
	(647)	(555)	(526)	(535)
Nearest Neighbor (1:1)		-2,312**	-1,570	-1,482
		(988)	(995)	(990)
Radius 1 (caliper=0.1)		-2,015***	-2,143***	-1,996***
		(520)	(540)	(522)
Radius 2 (regression adjustment)		-2,213***	-1,908**	-1,617*
		(827)		(834) (841)
Inverse Probability Weighting		-2,105***	-1,787***	-1,874***
		(550)	(523)	(528)
Kernel 1 (bw=0.02)		-2,118***	-1,879***	-1,840***
		(552)	(535)	(548)
Kernel 2 (bw=0.06)		-2,129***	-1,905***	-1,886***
		(548)	(527)	(531)
Kernel 3 (bw=0.2)		-2,198***	-2,161***	-2,097***
		(534)	(519)	(515)
No. of observations	6,737	6,737	6,737	6,737
Propensity score specification				
Personality traits			✓	✓
Inter-generational variables			✓	✓
Social network			✓	✓
Labor market flexibility			✓	✓
Life satisfaction			✓	✓
ALMP expectations			✓	✓
Socio-demographic characteristics		✓	✓	✓
Labor market history		✓		✓

Note: Depicted are estimated average treatment effects on the treated as the difference in mean outcomes between participants and matched non-participants using, OLS regression, inverse probability weighting (IPW), one-to-one nearest neighbor matching, radius with a caliper of 0.1, respectively regression adjustment (see Huber, Lechner, and Steinmayr, 2015) and epanechnikov kernel propensity score matching with bandwidths 0.02, 0.06 and 0.2. Standard errors are in parentheses and based on bootstrapping with 999 replications. ***/**/* indicate statistically significance at the 1/5/10%-level.

Table A.6: Consequences for the Average Treatment Effects on the Treated (ATT) for Different Matching/Weighting Estimators - Long-term Training

	Unconditional raw difference	Propensity score specification		
		Standard (1)	Auxiliary (2)	Extended (3)
Regular employed 30 months after entry (t+30)				
Regression	0.006	-0.027	-0.010	-0.021
	(0.019)	(0.021)	(0.021)	(0.021)
Nearest Neighbor (1:1)		-0.001	-0.036	0.003
		(0.032)	(0.033)	(0.034)
Radius 1 (caliper=0.1)		-0.006	-0.020	-0.014
		(0.020)	(0.021)	(0.021)
Radius 2 (regression adjustment)		-0.008	-0.036	-0.0005
		(0.028)	(0.029)	(0.030)
Inverse Probability Weighting		-0.026	-0.010	-0.018
		(0.021)	(0.021)	(0.021)
Kernel 1 (bw=0.02)		-0.027	-0.009	-0.019
		(0.022)	(0.021)	(0.022)
Kernel 2 (bw=0.06)		-0.025	-0.009	-0.017
		(0.021)	(0.021)	(0.022)
Kernel 3 (bw=0.2)		-0.012	-0.002	-0.010
		(0.02)	(0.02)	(0.02)
Cumulated earnings in € up to t+30				
Regression	-3,876***	-5,763***	-5,226***	-5,627***
	(938)	(829)	(804)	(823)
Nearest Neighbor (1:1)		-4,357***	-5,807***	-4,638***
		(1,556)	(1,595)	(1,637)
Radius 1 (caliper=0.1)		-4,961***	-5,298***	-5336***
		(790)	(814)	(827)
Radius 2 (regression adjustment)		-4,939***	-5,700***	-4,560***
		(1,375)	(1,367)	(1,352)
Inverse Probability Weighting		-5,794***	-5,406***	-5,761***
		(844)	(815)	(846)
Kernel 1 (bw=0.02)		-5,788***	-5,282***	-5,610***
		(877)	(855)	(896)
Kernel 2 (bw=0.06)		-5,656***	-5,280***	-5,618***
		(845)	(823)	(859)
Kernel 3 (bw=0.2)		-4,826***	-4,566***	-4,989***
		(795)	(784)	(797)
No. of observations	5,717	5,717	5,717	5,717
Propensity score specification				
Personality traits			✓	✓
Inter-generational variables			✓	✓
Social network			✓	✓
Labor market flexibility			✓	✓
Life satisfaction			✓	✓
ALMP expectations			✓	✓
Socio-demographic characteristics		✓	✓	✓
Labor market history		✓		✓

Note: Depicted are estimated average treatment effects on the treated as the difference in mean outcomes between participants and matched non-participants using, OLS regression, inverse probability weighting (IPW), one-to-one nearest neighbor matching, radius with a caliper of 0.1, respectively regression adjustment (see Huber, Lechner, and Steinmayr, 2015) and epanechnikov kernel propensity score matching with bandwidths 0.02, 0.06 and 0.2. Standard errors are in parentheses and based on bootstrapping with 999 replications. ***/**/* indicate statistically significance at the 1/5/10%-level.

Table A.7: Consequences for the Average Treatment Effects on the Treated (ATT) for Different Matching/Weighting Estimators - Wage Subsidies

	Unconditional raw difference	Propensity score specification		
		Standard (1)	Auxiliary (2)	Extended (3)
Regular employed 30 months after entry (t+30)				
Regression	0.128*** (0.022)	0.092*** (0.022)	0.093*** (0.022)	0.085*** (0.022)
Nearest Neighbor (1:1)		0.113*** (0.037)	0.096** (0.038)	0.092** (0.04)
Radius 1 (caliper=0.1)		0.105*** (0.021)	0.111*** (0.021)	0.098*** (0.022)
Radius 2 (regression adjustment)		0.090*** (0.032)	0.104*** (0.032)	0.087*** (0.032)
Inverse Probability Weighting		0.096*** (0.022)	0.09*** (0.022)	0.082*** (0.022)
Kernel 1 (bw=0.02)		0.097*** (0.022)	0.087*** (0.023)	0.084*** (0.023)
Kernel 2 (bw=0.06)		0.100*** (0.022)	0.095*** (0.022)	0.087*** (0.023)
Kernel 3 (bw=0.2)		0.119*** (0.021)	0.114*** (0.021)	0.109*** (0.021)
Cumulated earnings in € up to t+30				
Regression	4,783*** (1,105)	3,386*** (833)	3,861*** (787)	3,670*** (810)
Nearest Neighbor (1:1)		4,450*** (1,634)	3,528** (1,769)	3,966** (1,703)
Radius 1 (caliper=0.1)		4,168*** (772)	4,129*** (801)	4024*** (800)
Radius 2 (regression adjustment)		3,708*** (1,409)	3,632** (1,477)	3,647*** (1,450)
Inverse Probability Weighting		3,460*** (838)	3,769*** (797)	3,522*** (831)
Kernel 1 (bw=0.02)		3,749*** (852)	3,677*** (844)	3,604*** (883)
Kernel 2 (bw=0.06)		3,723*** (830)	3,924*** (805)	3,691*** (839)
Kernel 3 (bw=0.2)		4,450*** (801)	4,359*** (771)	4,245*** (785)
No. of observations	5,476	5,476	5,476	5,476
Propensity score specification				
Personality traits			✓	✓
Inter-generational variables			✓	✓
Social network			✓	✓
Labor market flexibility			✓	✓
Life satisfaction			✓	✓
ALMP expectations			✓	✓
Socio-demographic characteristics		✓	✓	✓
Labor market history		✓		✓

Note: Depicted are estimated average treatment effects on the treated as the difference in mean outcomes between participants and matched non-participants using, OLS regression, inverse probability weighting (IPW), one-to-one nearest neighbor matching, radius with a caliper of 0.1, respectively regression adjustment (see Huber, Lechner, and Steinmayr, 2015) and epanechnikov kernel propensity score matching with bandwidths 0.02, 0.06 and 0.2. Standard errors are in parentheses and based on bootstrapping with 999 replications. ***/**/* indicate statistically significance at the 1/5/10%-level.