Subjective Expectations and the Effectiveness of Labor Market Policies

Robert Mahlstedt*

July 14, 2019

Abstract

The paper shows that human capital-intensive labor market training programs, relying on market-based assignment procedures, are more effective when participants expect the program participation ex ante. The causal effect is identified based on a difference-in-difference approach comparing participants in programs with different assignment procedures. Individuals who expect to participate ex ante have larger incentives to search for an appropriate program, which increases the match quality and the program effectiveness. A causal forest approach is utilized to identify heterogeneous effects of expected participation probabilities on the program effectiveness.

Keywords: Labor market policies, Expectations, Treatment Effects, Heterogeneity

JEL codes: J08, J68, D83, D84

*University of Copenhagen, Denmark; IZA Bonn, Germany; `robert.mahlstedt@econ.ku.dk`
Corresponding address: University of Copenhagen, Department of Economics, Øster Farimagsgade 5, 1353 København K, Denmark. Tel: +45 35 33 79 22

I thank Steffen Altmann, Marco Caliendo, Peter Kuhn, Steffen Künn, Alexander Sebald, Gerard van den Berg and participants at the 2016 Meeting of the European Society for Population Economics (ESPE), the 2016 Meeting of the European Association of Labour Economists (EALE), the 2016 Workshop of the Centre for Research in Active Labour Market Policy Effects (CAFE), the 2017 Spring Meeting of Young Economists and seminars at IZA Bonn, the University of Potsdam, the University of Copenhagen and the Melbourne Institute. I further thank Philipp Schmalen for excellent research support and the IZA Bonn, as well as the IAB Nuremberg for access to the IZA/IAB Linked Evaluation Dataset (for more information see Eberle *et al.*, 2017). I gratefully acknowledge financial support by the Danish Ministry of Employment and I am thankful to the University of Potsdam and the German Research Foundation (Grant: Caliendo, CA829/7-1) for financial support for my research visits during which part of this project has been conducted.

1 Introduction

In modern welfare states, market-based mechanisms become increasingly popular for the provision of governmental services (see e.g. Sol *et al.*, 2005). Although, a high degree of competition between service providers and more individual choices aim to increase the efficiency of public policies, it also implies that subjective perceptions and expectations about incentive structures, institutional settings or the state of the market influence an individual's reaction to a particular policy.[1] In the unemployment insurance system, activation policies, such as training or job search assistance, represent one of the major tools to promote the unemployed's reintegration into the labor market, but assignment procedures often involve a high degree of discretion, such as discretionary decisions by caseworkers, or require potential participants to provide effort. For instance, various countries, like the US or Germany, rely on voucher systems to assign individuals to training programs (see Hipp and Warner, 2008), which implies that potential participants can choose among different types of programs or program providers. Such a system aims to increase the match quality between the unemployed and the program, but it also implies that potential participants can increase the effectiveness of a treatment by investing effort to identify promising programs and to search for an appropriate provider. In this context, the unemployed's subjective expectations about the future program participation play a central role since they determine the perceived incentives to invest effort in order to increase the match quality and therefore also the effectiveness of a treatment.[2]

This paper provides first evidence with respect to the relevance of such a mechanism by investigating the relationship between subjective expectations of unemployed workers about a potential future enrollment in labor market training and the effectiveness of a realized program participation later during the unemployment spell. I exploit a unique combination of survey and administrative data for a sample of newly unemployed job seekers in Germany. The survey data include self-reported information on the respondents' expectations

[1] Previous examples show that subjective beliefs about the pension system affect savings and retirement decisions (see e.g. Chan and Stevens, 2004; Bottazzi *et al.*, 2006), labor supply is determined by the perceived incentive structure inherent in the tax and transfer systems (Chetty and Saez, 2013) and consumption expenditures are related to inflation expectations (D'Acunto *et al.*, 2015) that in turn depend on the assumed targets of the central bank (Christelis *et al.*, 2016).

[2] Here, match quality refers to the match between the participant and the program, assuming that there are heterogeneous types of programs differing in quality and content, while gathering the relevant information creates additional costs for a potential program participant.

about the likelihood to participate in a program in the near future (given that they remain unemployed), hereafter referred to as the *expected participation probability*. Moreover, the administrative data provide detailed information on the subsequently realized program participation, as well as labor market outcomes. I can show that participants in human capital-intensive training programs (long-term training) are substantially more likely to be employed (about 10 percentage points) in the long-run when they do expect the treatment ex ante compared to those participants who do not expect the treatment. This can be explained by the fact that participants in long-term training can freely choose a program provider, which implies that their expectations about the participation have a strong impact on the individual incentives to search for the most effective treatment. To investigate whether estimates effect indeed reflects a causal relationship, I apply a difference-in-difference approach that exploits institutional differences with respect to assignment procedures for different types of training programs. Therefore, I consider a control group of participants in an alternative program (short-term training), who cannot choose a provider and therefore have no incentives to adjust their pre-treatment behavior based on their expected participation probability. It is shown that the pre-treatment expectations of the control group are unrelated to the effectiveness of the intervention.

The validity of the empirical approach clearly depends on the assumption that the expected participation probability is similarly related to unobserved factors that might be connected to subsequent labor market outcomes of the participants in the two programs. Therefore, I conduct a variety of balancing tests taking into account different pre-treatment outcomes, like past employment probabilities, job search characteristics and related expectations about reemployment prospects and earnings before the actual program participation. For both types of participants, the expected participation probability is similarly connected to all pre-treatment outcomes but has a different impact on outcomes measured after the beginning of the treatment, which strongly supports the validity of the empirical model and the causal interpretation of the results.

A further analysis sheds light on the underlying effect mechanisms. First of all, in a second survey wave, a subset of program participants answered several questions related to the assignment and the quality of the program. Participants who expect the treatment ex ante are more engaged in the assignment process, have greater influence on the type of the

program and are more satisfied with the treatment ex post. I can also that the effect of the expected participation probability is particularly pronounced for program participants who generally have a high motivation to exert effort into the preparation of the treatment. Moreover, I adopt a recently developed method from the machine-learning literature, the so-called causal forest approach, to identify background characteristics that are associated with heterogeneous effects of the participant's pre-treatment expectations. It is shown that, in particular, the unemployed's initial job search effort before participating in a program, their labor market histories and their sense of control are strong determinants of the impact of the expected participation probability on the effectiveness of long-term training programs. This is in line with theoretical considerations, as it can be argued that the effects are strongest for those groups who (i) have the largest incentives to search for the most effective program or (ii) experience the strongest impact of the match quality on the program effectiveness. Finally, I show that the results are robust with respect to the presence of other potential mechanism such as differences with respect to the timing of the treatment and participants' expectations about the effect of the intervention.

The results have several important implications. First, the paper provides first direct evidence that program participants can influence the effectiveness of labor market policies depending on their subjective expectations. This provides a better understanding why certain policies might be not very effective (see e.g. Card *et al.*, 2010; Kluve, 2010; Card *et al.*, 2017, for an overview of international ALMP studies) and emphasizes the importance of active counseling by the employment agency.[3] When potential participants are free to choose a specific program or a program provider, information policies targeting their expectations about the participation or their possibilities to collect the relevant knowledge can be an effective tool to improve the success of an intervention. Moreover, the paper adds to the growing literature analyzing how subjective perceptions can lead to suboptimal behavior of unemployed workers, for instance, due to systematical biases in the perception of job finding probabilities (see Spinnewijn, 2015), present-biased time preferences (see DellaVigna and Paserman, 2005; Paserman, 2008), reference-dependence (see DellaVigna *et al.*, 2017) or their sense of control (see Caliendo *et al.*, 2015; McGee, 2015). In the context of labor

[3]This is related to several studies pointing out the importance of counseling unemployed workers (see e.g. Gorter and Kalb, 1996; Behaghel *et al.*, 2014; Altmann *et al.*, 2017), analyzing the impact of caseworkers on job finding chances (see e.g. Behncke *et al.*, 2010a,b), as well as their efficiency when assigning ALMP programs (Lechner and Smith, 2007).

market policies, subjective expectations can be linked to the presence of anticipation effects, which implies an adjustment of the job search behavior as reaction to the possibility of a future treatment.[4] However, my results show that, beside the adjustment of the job search strategy, a second dimension of the individual behavior is affected by the expectation about a future program participation, which becomes important if the individual actually enters the program. Both types of behavioral adjustments should be considered when modeling the effect mechanisms of a policy or evaluating its costs and benefits. The reminder of the paper is organized as follows. Section 2 discusses the institutional details, the relevant hypotheses and the data. Section 3 presents the empirical strategy and discussed the underlying assumptions, while Section 4 shows the estimation results. Finally, Section 5 concludes.

2 Conceptual Framework and Data

2.1 Training Programs in Germany

The German unemployment insurance system provides an ideal example to study the importance of subjective expectations for labor market programs that rely on fundamentally different assignment procedures and resulting consequences for their effectiveness. For the empirical analysis, I consider two types of labor market training programs. First, *long-term training* aims to improve occupational-specific skills of the participants in order facilitate the reintegration into the labor market. The program is frequently assigned to unemployed workers in Germany and generally represents a class of human capital-intensive training programs that are commonly used in many Western countries. Program participation requires a high level of participants' commitment since the program typically last from several months up to one year, while average program duration is about six months. The treatment can take place in classrooms or simulated workplaces, including e.g. training in computer assisted bookkeeping, operating specific construction machines, as well as specialized courses in specific legal fields, marketing or sales strategies. Previous studies found positive effects

[4] On the one hand, the threat of participating in a program that reduces the individual utility level provides incentives to increase the search effort in order to find a new job before the treatment can be realized (see Black *et al.*, 2003 for the US; Geerdsen, 2006; Geerdsen and Holm, 2007; Rosholm and Svarer, 2008 and Graversen and van Ours, 2008, 2011 for Denmark, as well as Carling and Larsson, 2005 and Hägglund, 2011 for Sweden). On the other hand, the presence of a program that is assumed to be beneficial would encourage individuals to reduce the search effort and wait out until the treatment can be realized (see van den Berg *et al.*, 2014a for the UK and Crépon *et al.*, 2014 for France).

only in the very long run (e.g. Fitzenberger *et al.*, 2008; Lechner *et al.*, 2011) or even partly negative effects on employment prospects (e.g. Lechner and Wunsch, 2008). In the short-run, these programs are expected to create a strong locking-in effect since participants typically reduce their job search effort during the treatment.

The assignment to the program is organized through a voucher system. Therefore, the caseworker does not choose a specific course for the unemployed but hands out a training voucher to the job seeker (see e.g. Rinne *et al.*, 2013). Caseworkers are instructed to grant a voucher only if the estimated probability that a job seeker will find employment immediately after finishing the program is at least 70%. The voucher defines the objective, the content and the maximum duration of the course, but the unemployed are allowed to find an appropriate provider for themselves, respectively to not redeem the voucher (see Bernhard and Kruppe, 2012; Doerr *et al.*, 2017). This is an important feature as it implies that potential participants can make an additional effort, e.g. by searching for an appropriate program provider, to increase the match quality and therefore also the effectiveness of the treatment.

As a second intervention, I consider *short-term training* measures (see e.g. Wolff and Jozwiak, 2007), which last from two days up to eight weeks. The content of these programs is less occupational-specific and the treatment typically aims to activate unemployed workers by providing skills that improve and facilitate job search or serve as a test of the job seeker's abilities. In contrast to long-term training, participants are directly assigned to a specific treatment by their caseworker. Moreover, participants only spend a short period in the treatment. This implies that they have only limited incentives to invest effort to increase the match quality since such an investment is unlikely to increase the effectiveness of the treatment.[5] Although, the program participation is not mandatory in general (neither for short- nor long-term training), specific obligations, e.g. an ALMP participation or job search requirements, might have been defined on an individual basis. However, empirical evidence indicates that this is only relevant for a very small fraction of individuals in the present case.[6]

[5]See Osikominu (2012) for a general comparison of short- and long-term training programs in Germany.

[6]Unemployed workers are obliged to sign so-called integration agreements *(Eingliederungsvereinbarung)* (see e.g. Jacobi and Kluve, 2007; van den Berg *et al.*, 2014b), which define the job seeker's obligations and the services that she receives by the employment agency, including search activities, as well as ALMP participation. However, only about 19% of the estimation sample have signed an integration agreement at the moment of the first interview, while in general less than half of the integration agreements specify participation in an ALMP program (see Schütz *et al.*, 2011).

2.2 Hypotheses

The presence of these two different assignment procedures allows me to analyze the connection between the subjective expectations of participants and the effectiveness of the program. Following van den Berg *et al.* (2009), Appendix A describes a theoretical job search model where unemployed workers face the possibility to participate in a training program in the future. Moreover, due to voucher system potential participants in long-term training can directly influence the effectiveness of the realized treatment. Since such a mechanism does not exist for short-term training, one can derive the following hypotheses:

Hypothesis 1. *Long-term training programs are more effective if participants expect the treatment ex ante, while the effectiveness of short-term training programs is independent of the participants' pre-treatment expectations.*

Intuitively, a job seeker has the possibility to participate in a training program in the future, while potential programs differ with respect to their quality and content, which implies that the individual can increase the match quality by searching for an appropriate program or program provider.[7] However, this effort investment creates additional costs and, in many situations, potential participants have to decide how much effort they want to invest, when they still face uncertainty whether they will actually participate or not. Since searching for the most effective program only pays off if the job seeker actually enters the treatment, the incentives of a potential participant crucially depend on the expectation about the likelihood to enter the treatment. The voucher system implies that this mechanism is particularly important for long-term training since the choice of the program provider can be directly related to the match quality and the program effectiveness. However, participants in short-term training are directly assigned to a specific treatment by the caseworker, and therefore have no (or only little) incentives to adjust the individual behavior based on their pre-treatment expectations. Moreover, the nature of the intervention, in particular the short program duration, implies that the match quality might be less relevant for the program effectiveness.

[7] Alternatively, one could also argue that program participants can directly increase the effectiveness of a given treatment by preparing themselves, e.g. by acquiring relevant materials or skills or adapting their job search behavior to be compatible with the treatment.

Hypothesis 2. *The impact of the expected participation probability on the program effectiveness is large if 1) individuals adjust their effort more strongly in response to their pre-treatment expectations or 2) the marginal impact of the match quality on the program effectiveness is large.*

Since the expected participation probability is assumed the influence the program effectiveness through the participant's pre-treatment effort, there are two mechanisms that could create heterogeneous effects. First, there could be heterogeneity in the way that individuals adjust their effort in response to the expected participation probability. For instance, there could be differences with respect to time constraints or expected relevance that could prevent some individuals from searching for a program or program provider, although they expect to participate in the future. Second, the match quality between the participant and the provider could be more important for some individuals than for others, which would also imply heterogeneity with respect to the impact of the expected participation probability on the program effectiveness. The relevance of different factors that are potentially associated with these two mechanisms are discussed in Section 4.3.

2.3 Data

The empirical analysis exploits the *IZA/IAB Linked Evaluation Dataset*, which includes survey information on individuals who entered unemployment between June 2007 and May 2008 in Germany (see Caliendo *et al.*, 2011; Eberle *et al.*, 2017). About 17,400 individuals are interviewed shortly after the entry into unemployment (between 7 and 14 weeks). Besides an extensive set of individual-level background characteristics (including socio-demographic and household information, as well as personality measure), the respondents are asked a variety of non-standard questions about their subjective assessments on future economic outcomes and job search characteristics. Most importantly, this includes expectations about the participation in labor market programs (see Section 2.4 for details), the job search behavior, expectations about future earnings and employment prospects, as well as measures that describe the role of the employment agency in the job finding process. For the 88% of individuals who agreed, these survey data are merged to administrative information from the *Integrated Employment Biographies* (IEB) provided by the Institute for Employment

Research (IAB).[8] The IEB integrates different sources, e.g., employment history, benefit recipient history, training participation history and job search history and therefore provides detailed information on labor market histories, as well as outcomes such as employment states, earnings, transfer payments and participation in active labor market policies for a period of 30 months after the entry into unemployment. Altogether, this amounts to a total of 15,173 realized interviews. A subset of respondents (about one fourth) is also asked additional questions including their motivation to contact the employment agency, while about 55% of the initial sample is interviewed again one year after the first interview. Beside a repetition of the initial survey, at the second interview participants in long-term training are also asked a variety of questions assessing the assignment to and the quality of the prior treatment. These additional information are exploited in Section 4.2.

The combination of survey and administrative data provides an ideal setting to analyze empirically the mechanisms discussed before since the dataset provides measures of expectations about a future participation in training programs obtained at the beginning of the unemployment spell and information about the subsequently realized program participation. For the purpose of the study, the estimation sample is restricted to all individuals who remain unemployed and do not participate in any labor market program between the entry into unemployment and the first interview and report non-missing information for the relevant expectation measures discussed below. Individuals are defined as participants if they attend a program within the first 12 months of the unemployment spell. Based on this definition, the estimation sample includes 707 participants in long-term training 1,457 participants in short-term training programs and 4,301 individuals who do not participate in any training program (see Figure B.1 in the Appendix for a graphical illustration of the empirical setting).

2.4 Measuring Expectations

The key variable for the analysis is given by the subjectively expected probability that the respondent will participate in a long-term training program in the near future, which will be also denoted as the expected participation probability π in the following. This information is measured by the question: *"Assuming that you are still unemployed during the next three*

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

months. What do you think is the probability that you will participate in a long-term training scheme within this period?" Possible answers range from zero (very unlikely) up to ten (very likely). It is important to note that the respondents are asked to give their assessment conditional on not finding a job, which implies that differences with respect to the expected reemployment probability should not influence the respondents' answers.

The distribution of this variable, separated by the actual treatment status, is depicted in Figure 1. In general, most individuals report the answers zero, five or ten, while there is a correlation between the expected and the actual treatment status. For example, about 35% of the participants in long-term training report ex ante that it is very likely that they will participate, while only about 14% of the non-participants do. In line with this, about 30% of the non-participants say ex ante that is very unlikely that they will participate, while only 17% of the participants report a zero. Based on this information about the expected participation probability, I construct a binary measure, which divides the sample into a group who expects the treatment summarizing answers 5-10 (π-high) and a group who does not expect the treatment containing answers 0-4 (π-low) (see van den Berg *et al.*, 2009, who use the same variable without exploiting information on the actual participation decision).

[Insert Figure 1 about here]

A second type of expectation measure refers to the expected treatment effect, which will be also denoted as δ in the following and is covered by the survey question: *"In your opinion, to what extent would your chances of finding new employment be changed by participating in long-term training?"* The answers are measured on a 5-point scale ranging from 'improve strongly' to 'worsen strongly' and can be interpreted as a proxy for the expected differences in job finding rates between the treated and non-treated situation. The distribution of this variable separated by the realized treatment status is shown in Figure B.2 and the joint distribution of both types of expectations is depicted in Table B.1 both presented in the Appendix. The interdependence of the expected treatment effect and the expected participation probability is analyzed in Section 4.4.

2.5 Observed Differences in Labor Market Outcomes

Table 1 shows differences with respect to the main outcome variables between individuals who expect the program participation ex ante (π-high) and those who do not (π-low). In particular, I focus on the employment status 12 months, respectively 30 months after the entry into unemployment, which characterizes the end of the observation period. The outcome variables are measured relative to the entry into unemployment and not the start of the treatment since the timing of the program start can be causally affected by the expected participation probability. This is further discussed in Section 4.4.[9] Moreover, I also consider the average monthly earnings given that the individual was employed in the corresponding month within the full observation period of 30 months. Panel A shows that participants in long-term training who expect the treatment ex ante have substantially higher employment rates than those who do not expect the treatment. The difference between the two groups of program participants is about 13 percentage points, statistically significant at the 1%-level and relatively constant over time, while there are no significant differences with respect to average earnings. Although it does not take into account that participants who expect the treatment ex ante might be fundamentally different from those who do not expect the treatment, it documents that the effectiveness of long-term training programs is connected to the participants' expected participation probability. The aim of the following analysis is to investigate whether this relationship can be attributed to the mechanism discussed before and reflects a causal effect, which could potentially justify policy interventions.

[Insert Table 1 about here

As a first step, the sample of participants in short-term training is also divided with respect to the expected participation probability π.[10] As shown in Panel B, for participants in short-term training, expecting a treatment ex ante is associated with a six percentage point lower employment probability one year after the entry into unemployment, while there is no difference in employment rates in the long-run between those participants who expect the treatment and those who did not. Moreover, as shown in Panel C, there are only

[9] The results are qualitatively similar when measuring the outcome variables relative to the program start.

[10] Note that here π also refers to the expected participation probability in long-term training, which allows a direct comparison with the results from Panel A. As a sensitivity analysis, I also consider an alternative measure that covers subjective expectations about the general participation in any labor market program (see Table B.3).

minor differences within the group of non-participants. When comparing participants to non-participants, it can be seen that the employment rates of participants in long-term training are generally lower 12 months after the entry, which is not surprising given that the treatment is typically associated with a strong locking-in effect. However, at the end of the observation period participants in long-term training who expect the treatment ex ante have the highest employment rate among the six groups, while participants who do not expect the treatment have substantially lower employment rates than all other groups. The descriptive comparison of the three groups based on their realized treatment status provides first evidence that the strong positive connection between the long-run employment rates and the expected participation probability is very program-specific, as it only appears for participants in long-term training. This indicates that it cannot be explained by other characteristics that be might correlated with subjective expectations and the employment prospects in general.

3 Empirical Strategy

3.1 Difference-in-Difference Model

In the following, the idea of comparing the differential effect of the expected participation probability π on the effectiveness of the two programs is formalized in a difference-in-difference (DID) model. As discussed in Section 2.2, individuals who expect to participate ex ante have higher incentives to search for an appropriate program to increase the match quality and therefore also the effectiveness of the treatment. Since the assignment to long-term training is organized through the voucher system, such a mechanism is assumed to be particularly important for long-term training, while it plays no role for short-term training programs where the treatment is directly assigned by the caseworker. These institutional differences with respect to the assignment process provide an ideal basis to estimate the causal effect as the differential impact of the expected participation probability for the two programs. The model is characterized by the following equation:

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 \pi_i^{high} + \beta_3 (D_i \times \pi_i^{high}) + v_i, \tag{1}$$

where Y_i denotes the outcome variable of interest, D_i indicates a dummy that takes the value one for participants in long-term training and zero for individuals participating in short-

term training, while π_i^{high} indicates an expected assignment probability of five of higher. The coefficient β_3 related to the interaction term between the treatment status and the expected participation probability π then indicates the differential impact of π on the effectiveness of long-term relative to short-term training.

To allow the DID model to have a causal interpretation it is required that the expected participation probability π is similarly related to unobserved characteristics (that are relevant for labor market outcomes) for both groups of participants. If this is this case, the DID approach is suitable to control for unobserved heterogeneity, i.e. factors that are related to the pre-treatment expectations and subsequent labor market outcomes simultaneously. To test the validity of this assumption, I conduct different types of balancing and placebo tests that will be discussed in detail in the following section and provide strong support that the underlying identification assumption is fulfilled. Moreover, I condition on a rich set of control variables such as socio-demographic and household characteristics, regional and seasonal information, personality traits, i.e. the 'Big Five' factors and locus of control, as well as detailed labor market histories, including characteristics of the previous job and (un)employment experiences within the last ten years before the entry into unemployment, which has been shown to be particularly important in the context of training programs (Mueser *et al.*, 2007; Lechner and Wunsch, 2013; Biewen *et al.*, 2014; Caliendo *et al.*, 2017). To estimate β_3 based on these control variables, I use inverse probability weighting (IPW) with weights obtained from probit estimations.

3.2 Selection into the Treatment and Balancing Tests

The following section shows various tests that support the validity of the identification assumption. First of all, Table 2 shows the marginal effects of probit models estimating the impact of the expected participation probability π on the actual probability to participate in the different programs. It can be seen that reporting an expected participation probability π of 5 or higher is associated with an increased likelihood to actually participate of 8-9 percentage points. As shown in column 1 and 2, the coefficients related to the two different programs (short-term and long-term training) are very similar, which provides a first indication that the expected participation probability is similarly related to the selection into the two different programs. Moreover, to provide direct evidence for the validity of the un-

derlying assumption of the DID model, I conduct balancing tests for all observed covariates. Therefore, I estimate the DID model specified in Equation (1) considering the background characteristics as outcome variables. As shown in Table 3, even in the unconditional specification only four out of 53 coefficients are statistically significant at the 5%-level, which indicates that the expected participation probability is similarly related to both types of participants. When applying IPW, no statistically significant differences can be found.

[Insert Table 2 and Table 3 about here]

In a next step, I conduct a placebo test, which considers different measures of the individual job search behavior and related expectations about realized labor market outcomes, i.e. expected reemployment prospects, earnings and the impact of the employment agency on job finding,[11] as the outcome of Equation (1). Importantly, all these characteristics are obtained before the individual enters the program and therefore none of the variables can be affected by the causal mechanism discussed in Section 2.2. However, the variables are particularly informative about the presence of potential unobserved heterogeneity since they characterize the main components of the job search process and are important determinants of the selection into a particular treatment. Moreover, the assumed influence of the employment agency on the job finding prospects provides a proxy for the (assumed) quality of the caseworker. In absence of unobserved heterogeneity, the expected participation probability π should have a similar impact on the corresponding variables for both types of participants, while any difference would suggest that there exist unobserved factors and would challenge the identifying assumptions of the DID. The results are summarized in 4. Interestingly, the expected participation probability π is unrelated to the job search behavior, i.e. the log average weekly number of applications and the number of search channels utilized, for both groups of program participants, which leads also to small and insignificant coefficients for the DID model presented in column 3. When considering the expected labor market outcomes, it can be seen that the expected job finding prospects and the assumed influence of the employment agency are indeed related to π. However, the correlation is very similar for both types of program participants and therefore the resulting estimates of the DID model are

[11] The expected reemployment probability is measured on a four-digit scale from 'very low' to 'very high'. The joint distribution of the expected reemployment probability, the expected income and the expected influence of the employment agency for both groups of participants are shown in Figure A.1 in Appendix A.

again small and statistically insignificant. In summary, the findings provide strong evidence that the underlying identification assumption of DID approach is fulfilled.

[Insert Table 4 and Figure 2 about here]

Finally, Figure 2 graphically illustrates the difference-in-difference comparison between participants in short- and long-term training based on the expected participation probability π for employment rates over a period from ten years before the entry into to unemployment up to the end of the observation period 30 months after the entry. It can be seen that π has no significant impact on the differences in the average yearly employment rates before the entry into unemployment and there is no particular trend within the ten-year period before the entry into unemployment. This indicates that time-constant unobserved characteristics, e.g. different levels of intrinsic motivation or ability, associated with the expected participation probability π are similar for both groups of participants and there is no evidence for the presence of time-varying unobserved factors. The latter is particularly important as it shows that the common-trend assumption is fulfilled, which allows me to also estimate a difference-in-difference-in-difference (DDD) model. Therefore, I consider ΔY_i as the outcome variable, which refers to the difference between the outcome variable of interest, e.g. the employment rate at a particular point in time, and the corresponding pre-treatment outcome measured before the entry into unemployment, e.g. average employment rate in a given reference period.

3.3 Conditional Average Treatment Effects

In a subsequent analysis, I aim to identify heterogeneous effects of π on the program effectiveness across different subgroups. As outline in Section 2.2, the magnitude of the effect depends on the extent of behavioral change in response to the expected participation probability, respectively the relevance of the match quality (between the participant and the program) for the subsequent employment outcomes, e.g. whether the choice of the program provider is important for the success of the treatment. To identify the relevant heterogeneities in a principled way, I adopt a recently developed approach from the machine learning literature, which is based on the concept of regression trees and random forest algorithms (see Breiman,

2001).[12] The aim of the so-called causal forest (CF) algorithm (see Athey and Imbens, 2016; Wager and Athey, 2018) is to identify those characteristics that are particularly important when estimating conditional average treatment effects (CATEs) of the expected participation probability π in the sample of participants in long-term training:

$$E[Y(\pi^{high}) - Y(\pi^{low})|X = x, D = 1], \tag{2}$$

The CATEs can be interpreted as the group-specific effects on individuals that share similar characteristics X. To identify groups that are similar, a tree-based algorithm forms disjoint groups of observations, called leaves, which share values of certain Xs. Starting with the full dataset, for each $X_j = x$, the algorithm forms a candidate split into two leaves. An observation with $X_j \leq x$ is placed in the left leave, while an observation with $X_j > x$ follows the right leave. From all candidate splits, the one that minimizes a goodness-of-fit criterion is then implemented.[13]

Predicting heterogeneous effects of π based on the CF approach allows me to identify subgroups who show a particularly strong response with respect to their pre-treatment expectations. This information can be used to examine the validity of the proposed mechanism and the resulting policy implications. Moreover, in the present empirical setting, the findings can be also used to verify the validity of the empirical approach discussed in Section 3.1. The CF identifies heterogeneous effects of π in the sample of participants in long-term training who have (the strongest) incentives to react to their pre-treatment expectations about the participation probability. The generated algorithm is then used to predict CATEs for the control group of participants in short-term training (or non-participants) who are assumed to have no incentives to adjust their behavior based on their pre-treatment expectations. If the latter is true and there are no unobserved confounders associated with π and the labor market outcome Y, the predicted CATEs (generated based on participants in long-term

[12]Given the large set of available background characteristics, who are potential candidates for analyzing heterogeneous effects, a data-driven feature selection reduces the risk of mistaking noise for a true treatment effect or ignoring unanticipated results.

[13]In contrast to regression trees, standard fit measures like the mean squared error (MSE) are not feasible when considering CATEs since $Y(\pi^{high}) - Y(\pi^{low})$ is not observed at the individual-level. However, as shown by Athey and Imbens (2016) minimizing the expected MSE of predicted treatment effects can solve this problem by dividing the sample into different subsample. One subsample is then used to determine the splits in the tree, while another subsample is used to estimate $\hat{\tau}_l = \bar{y}_l^{high} - \bar{y}_l^{low}$, which denotes the CATE within a certain leave of the tree l, and calculate the corresponding MSE. Wager and Athey (2018) extend this approach to many trees in a causal forest, by using predictions from a large number of trees based on subsamples of the entire dataset.

training) should be independent of the realized effect of π in the other groups (out-of- sample predictions). A positive correlation, however, would show that there are common factors that are similarly related to the impact of π in both groups, which could be interpreted as an indication that the DID approach might suffer from endogeneity bias.[14]

4 Estimation Results

This section shows empirically how the expected participation probability influences the subsequent labor market outcomes of participants in training programs. The first part of the analysis presents the average effect of the expected participation π based on the DID model discussed in Section 3.1, while the subsequent analyses shed light on the underlying effect mechanism by exploiting more detailed information about assignment into the treatment and analyzing heterogeneous effects based the CF approach. Finally, I investigate the sensitivity of the results with respect to unobserved heterogeneity and other potential mechanisms.

4.1 Subjective Expectations and the Program Effectiveness

Table 5 shows the average effect of reporting an expected participation probability of five or higher (π-high) relative to individuals who report an expected participation probability lower than five for the three labor market outcomes of interest. The estimated effects are separated based on the realized treatment status within the first 12 months of the unemployment spell. The first column shows the results for participants in long-term training. There is a strong positive effect of expecting a treatment on the employment probabilities after the individuals actually entered the program. Twelve months after the entry into unemployment, those participants who expected the treatment ex ante are 12.1 percentage points more likely to be regular employed relative to those who do not expect to participate. The effect is statistically significant at the 1%-level. Although the effect is slightly lower (10.3 percentage points) 30 months after the entry, expecting the treatment ex ante also has a strong positive influence on the long-run effectiveness of the program, which is statistically significant at the 5%-level. There is no significant effect on average monthly earnings. The second column shows the corresponding effect of the expected participation probability π on participants

[14]The results of this test will be discussed in Section 4.4.

in short-term training. In contrast to long-term training, π has a negative effect on the employment probability after 12 months of about 6.8 percentage points, no effect on the employment probability after 30 months and also no statistically significant effect on the average monthly earnings. Moreover, column (3) shows the effect of π for individuals who do not participate in a training program. The effects on non-participants are small and statistically insignificant for all outcome variables.

[Insert Table 5 about here]

Completing the graphical presentation of Figure 2, column (4) to (6) show the results of the DID model estimating the differential impact of the expected participation probability π on participants in long-term training relative to participants in short-term training. At the end of the observation period (30 months after the entry into unemployment), there is a positive employment effect of the expected participation probability π on participants in long-term training relative to participants in short-term training of about of about 10.3 percentage points, which is statistically significant at the 10%-level. Moreover, columns (5) and (6) show the findings for conditional DDD models, which additionally account for two levels of baseline employment rates in the last two (column 5), respectively five (column 6) years before the individual became unemployed. Again, there are positive employment effects of the expected participation probability, which are statistically significant and similar (or even larger) compared to the baseline results. Again, there is no significant effect on average earnings, where I exploit previous income before unemployment as the reference level for the DDD model.

In summary, the baseline results confirm Hypothesis 1 since there is a positive and program-specific effect of the expected participation probability π on the labor market outcomes of participants. For long-term training programs, where participants have strong incentives to search for a program provider to increase the match quality, expecting the participation ex ante leads to a higher program effectiveness, while there is no effect for participants in short-term training, who have little incentives to adjust their pre-treatment behavior based on their expectations. This implies that the positive effect of the expected participation probability π is unlikely to be the consequence of endogenous factors associated with the participants pre-treatment expectations. In particular, the higher employment rates of participants who expect the treatment ex ante cannot be traced back to the fact

that a high expected participation probability is generally related to more favorable unobserved characteristics, since otherwise there would be also a positive effect on participants in short-term training (and non-participants). As already discussed in Section 3.1, the validity of this strategy (comparing the effect of π on different types of treatments) is supported by the fact that the expected participation probability π is similarly related to pre-treatment outcomes, i.e. past employment rates, related expectations measured at the beginning of the unemployment spell and other individual-level characteristics, for both types of participants.

4.2 Pre-Treatment Behavior and Match Quality

To provide further evidence for the validity of underlying mechanism, I exploit additional information provided by the survey data. First of all, in the second wave of the survey (about one year after the entry into unemployment) a subsample of individuals who participated in a long-term training program is asked a variety of question related to the prior treatment. This includes an assessment of the program quality, but also questions related to the pre-treatment behavior that could be interpreted as proxies for the individual effort devoted to the program preparation.[15] As shown in Panel A of Table 6, there are significant differences with respect to the effort related to the program preparation between those who expect the treatment ex ante and those who did not. In particular, those participants who expect the treatment ex ante report less often that the initiative to participate stemmed from the caseworker (9% vs. 14%), while the treatment more often has been a joint initiative of the caseworker and the unemployed (33% vs. 23%). Moreover, those who report a high expected participation probability had more often influence on the type (48% vs. 38%) and the timing (31% vs. 22%) of the program, while they have been also more likely to utilize their social network for the placement into the training program (14% vs. 8%). The results provide evidence that participants in long-term training who expect the treatment ex ante indeed devote more effort to preparation of the treatment. Moreover, as shown in Panel B, the expected participation probability is also connected with the ex post assessment of the program quality. Participants who expect the treatment report a significantly higher satisfaction with the program and assume that the employment prospects improved more

[15] Due to panel attrition, the sample interviewed in the second wave comprises only a subsample of all participants in long-term training.

strongly due to the treatment, which indicates a higher match quality. Finally, there are no differences with respect to the likelihood of completing the program and receiving a degree (see Panel C).

[Insert Table 6 about here]

Moreover, a subset of the original sample (about one fourth) is asked in the first survey wave about the motivation to contact the employment agency.[16] One of the possible answers refers to the intention to participate in a training program, which is particularly interesting since those who contact the employment agency in order to participate also seem to be more likely to actively search for the most effective program. Unfortunately, the sample of participants in long-term training who answered questions about their motivation is too small to directly utilize this information. Therefore, I consider a sample of 1,044 non-participants who report their motivation and estimate an index that summarizes the general probability to contact the employment agency to participate in a training program based on the full set of covariates. Afterwards, I divide the full sample of participants in long-term training based on the values of this generated index and estimate separate effects of the expected participation probability π on the labor market outcomes of program participants with a low (below median), respectively high (above median) probability to contact the employment agency to participate in a training program.

[Insert Table 7 about here]

As shown in Table 7, the impact of π is substantially stronger for participants who are assumed to have a high motivation to search for an appropriate program and increase the match quality. For instance, expecting to participate in the program ex ante increases the employment probability 30 months after the entry into unemployment by 17 percentage points (statistically significant the 5%-level) if the participant is assumed to have a high level of motivation. However, for those who are assumed to have a low level motivation the effect is only about six percentage points and statistically insignificant at conventional levels. The results are particularly interesting as they support the interpretation that the

[16]For the initial survey wave, 12 monthly cohorts are interviewed between June 2007 and May 2008. Three out of those 12 cohorts are asked a variety of additional questions including the motivation to contact the employment agency. This refers to 1,625 individuals out of the estimation sample including non-participants and participants in short-term and long-term training.

pre-treatment expectations influence the program effectiveness through the pre-treatment behavior of potential participants. Those individuals who are generally assumed to have a higher willingness to exert effort in connection with the preparation of a treatment seem to react more strongly to the expected participation probability. It should be also noted that there is a positive and statistically significant impact on the average monthly earnings for those who are assumed to have a high level of motivation, which further supports the idea that the expected participation probability is connected to the match quality. In summary, the additional survey information provide strong evidence that participants in long-term training who expect to participate ex ante are more likely to invest effort associated with the preparation of the treatment and that this effort investment is responsible for the positive effect of the expected participation probability on the effectiveness of the treatment.

4.3 Effect Heterogeneity

In the next step, I investigate heterogeneous effects of π based on observed background characteristics to provide further evidence with respect to underlying mechanism. As discussed in Section 2.2, the effect of π should be the strongest for individuals who show a strong behavioral response to the expected participation probability, respectively for those where the marginal impact individual behavior on the program effectiveness is large. Therefore, I exploit the CF approach presented in Section 3.3 to identify background characteristics that are particularly relevant when predicting the effect heterogeneity. The relevance of a variable in a CF is characterized by its feature importance, which is given by the average reduction in the MSE after including the specific variable (see Breiman, 2001). Following Athey and Wager (2019), I first estimate a pilot CF including all potential background characteristics. This comprises 33 features in total, which are categorized into six groups: socio-demographic characteristics, labor market histories, regional/seasonal information, personality traits, job search characteristics and related expectation measures.[17] In a second step, I train another CF including only the most important features within each of the six categories. This allows to eliminate confounding factors across variables that are likely to be correlated and identify those features that are truly relevant for predicting heterogeneous effects of π.[18] As the out-

[17]The feature importance for the full set of variables can be obtained from Figure B.3 in the Appendix.
[18]In particular, the second stage includes only variable with a feature importance above the group average.

come variable characterizing the program effectiveness, I consider the employment status at the end of the observation period (30 months after the entry into unemployment).

In general, the estimated effects of π vary between 4 and 21 percentage points (see Figure B.4 for the distribution of CATEs) suggesting a large heterogeneity with respect to the effect of the expected participation probability on the effectiveness of long-term training.[19] Moreover, Figure 3 shows the feature importance of the final specification. It can be seen that the labor market history, in particular the individual's unemployment experience and the last wage, explain the largest share of the variation with respect to the estimated effects. Additionally, the participants personality, especially the locus of control, and characteristics of the job search process, i.e. the number of own job applications, turn out to be important predictors of the effect of the expectation measure. Figure 4 shows the relationship between the estimated CATEs and the four most important features. It can be seen that for all four features, the estimated effects seem to decrease with increasing values of the corresponding variable for the areas with the majority of observations.

[Insert Figure 3 and Figure 4 about here]

Moreover, Figure 5 shows the heterogeneity of the average treatment effects with respect to the four most important features identified by the CF. Therefore, the sample of participants in long-term training is divided at the median of the corresponding variable and separated effects of the expected participation probability π on the employment status 30 months after the entry into unemployment are estimated. In the following, I discuss the estimated effects in light of Hypothesis 2 discussed Section 2.2.

Job search effort: Assuming that searching for a job or a program provider requires individuals to invest limited resources, there is a trade-off between the two types of effort. Those individuals who spend less effort on job search activities have capacities to devote more effort into the program preparation and therefore π would have a stronger impact on the program effectiveness. As discussed theoretically in Appendix A, the search effort (here the average weekly number of job applications) might be already affected by the subjective expectations about the future program participation and a subgroup analysis based on this

[19] It implies that the variation in treatment effects is more than one and half times the size of the average treatment effect of π and about 35% of the average of those participants not expecting the treatment ex ante.

endogenous variable would lead to biased results. Therefore, I first estimate a summary index of the search effort and classify individuals based on their score on this index. To ensure that the sample of interest (participants in long-term training) does not contribute to the development of the index, I use only data from the control group (participants in short-term training) to estimate the determinants of the overall search effort taking into account all other variables in the final specification of the CF. I then use this model to predict the expected level of search effort for each participant in long-term training. As shown in Panel A of Figure 5, there is a strong positive effect of the expectation participation probability on those individuals with low expected effort levels (17.6 percentage points), which is statistically significant at the 1% level, while the effect on those with high expected levels of search effort is close to zero and statistically insignificant. The findings are line with the theoretical predictions as the positive effect of the expected participation probability appears for those individuals who are assumed to have the capacity to search for a program or program provider.

Locus of control: The locus of control is a psychological concept that describes the degree to which people believe that they have control over the outcome of their life (see Rotter, 1966). Individuals with an internal locus (sense) of control believe that life's outcomes are due to their own efforts, while those with an external locus of control believe that outcomes are due to external factors (see Gatz and Karel, 1993).[20] It implies that program participants with an internal locus of control generally have higher incentives to spend effort into the treatment, e.g. by searching for a program or cooperating during the treatment. Assuming that the marginal returns with respect to the effort devoted to the preparation of the intervention are decreasing in these two dimensions,[21] the impact of the expected participation probability π would be stronger for individuals with an external locus of control since an additional unit of preparation effort has a larger impact on the program effectiveness. As shown in Panel B, the effect of the expected participation probability π indeed only appears for participants with an external locus of control. Expecting the treatment ex ante

[20]Several previous studies show that an internal locus of control is associted with significantly higher earnings (see Andrisani, 1977; Heineck and Anger, 2010; Semykina and Linz, 2007) and higher levels of search effort (see McGee, 2015; Caliendo *et al.*, 2015).

[21]This can be justified by two arguments. First, the quality of the program itself can be assumed to be more important for individuals with little intrinsic motivation. Second, the marginal returns of searching for a program provider should be decreasing with respect to effort devoted to the provider search.

increases the employment rate by about 21.6 percentage points for those with an external locus control, while the effect is close to zero for those with in internal locus of control.

[Insert Figure 5 about here]

Labor market history: Moreover, the individual's labor market history is identified as a strong predictor of the effect heterogeneity. First, it can be seen that the effect of π appears only in the sample of participants with a previous wage below the median (see Panel C). In this group, expecting the treatment ex ante leads to a higher employment rate of 20 percentage points (the effect is statistically significant at the 1%-level), while there is no effect on participants who had a previous wage above the median. This pattern could be explained assuming that the last wage is an indicator for the participants ability. It seems plausible that the potential of the treatment to improve the employment prospects decreases with the participants initial abilities, which implies that the program quality has a stronger impact on the reemployment prospect of participants with a low level of abilities. If this is true, the match quality and therefore also the expected participation probability π are more important for participants who earned a low wage in the previous job. Moreover, job seekers with high abilities (who earned a high wage before) can be generally assumed to have a shorter unemployment duration and therefore the unconditional likelihood of participating in a training program is substantial lower compared to those with lower abilities.[22] This implies that potential participants with high abilities might actually have lower incentives to search for a provider as they might assume to leave unemployment before the treatment could be realized. As a second variable characterizing the labor market history, the previous unemployment experience is a predictor of the effect heterogeneity. The results show that the expected participation probability π affects the program effectiveness only for individuals who have relatively little unemployment experience. This can be rationalized assuming that individuals who spent a lot of time in unemployment already gathered information about potential treatments in the past and therefore more often rely on these prior experiences when deciding about a specific program or provider. Those who have only little experience are required to invest more effort into the search for the most effective program and therefore the subjective expectations about the participation probability have a greater impact.

[22]Note the expected participation probability π denotes the conditional likelihood to enter the treatment assuming that the individual remains unemployed for another three months.

4.4 Sensitivity Analysis

Comparing predicted and realized effects of π: To test the validity of the empirical model, I use the predicted CATEs based on the causal forest algorithm to divide the sample into a subgroup that is predicted to respond strongly to the expected participation probability π and one that is not. As discussed in Section 3.3, the causal forest identifies the predicted effect of π if the individual would participate in a long-term training program. Hence, assuming that there are no unobserved confounders, the predictions should be unrelated to the impact of π on labor market outcomes for individuals who do not participate in long-term training. As shown in column 1, the CF approach identifies distinct effects within the sample of participants in long-term training. Across those with a predicted CATE below the median, there is negative and significant effect of 21.3 percentage points, while there is a positive effect of 44.4 percentage points in the sample of individuals with a predicted effect above the median. When considering the out-of-sample predictions for participants in short-term training (column 2) and non-participants (column 3), there is again no evidence that the expected participation probability π has any effect for individuals in those groups. Moreover, the factors that determine the impact of the expected participation probability for participants in long-term training are not informative to identify distinct effects in the other two groups. This supports the notion that there is a causal effect of π, which is program-specific and cannot be explained by unobserved characteristics.

[Insert Table 8 about here]

Timing of the program start: So far, the outcome variables have been measured relative to the moment when the individual enters unemployment not taking into account the timing of the treatment. This strategy is motivated by the fact that any differences in the duration between the entry into unemployment and beginning of the program are likely to be affected by the expected participation probability π. For instance, if individuals who expect the treatment ex ante prepare themselves more intensively, this could lead to an earlier start of the program. Descriptive evidence shows that this might be the case since participants in long-term training who expect the treatment ex ante start on average 1.5 months earlier than those participants who do not expect the treatment. As this could be a causal consequence

of the participant's expectations, conditioning on the timing of the program start would underestimate the actual effect of the expected participation probability π.

[Insert Figure 6 about here]

However, an alternative explanation could also imply that potential program participants already have private information about the timing of the program start and therefore adjust their expectations accordingly. These differences could potentially also translate into differences in long-term employment rates between participants with low or high expected participation probabilities π even if both groups would behave completely identical otherwise. To test the empirical relevance of this channel, Figure 6 shows the impact of the expected participation probability π for participants on the cumulated exit rate from unemployment for the first 18 months after the program start. Figure 6a shows the unconditional impact of π for participants in long-term training, while Figure 6b presents the corresponding DID estimates based on IPW. It can be seen that the expected participation probability has a positive and significant effect on the likelihood that a participant in long-term training leaves unemployment, which is particularly pronounced nine months after the program start. The magnitude of the effect is similar compared to the baseline results, which implies that π indeed has an impact on the program effectiveness beyond the mechanical effect that would be induced by delayed treatment starts. However, it should be noted that the results presented in Figure 6 most likely characterize a lower bound if the expected participation probability causally reduces the duration between the entry into unemployment and the program start.

Relevance of expected treatment effects: There is a positive correlation between the expected probability to participate in a training program and the expected effect of the treatment (see Table B.1). Therefore, it might be the case that the positive effect of π reflects the fact that program participants have private information about the effectiveness of the treatment in advance and therefore adjust their expectations about the likelihood to actually enter the treatment accordingly. For instance, those who expect the program to be beneficial could be more likely to assume that they find an appropriate provider, while those who expect the treatment to be inefficient might plan to not redeem the voucher. To test whether such a mechanism could explain the results, Table 9 shows the impact of the expected participation probability π conditioned on the expected effect of the treatment δ.

Interestingly, as shown in Panel A and B, the expected participation probability has a strong effect for those participants who expect the treatment to be not ($\delta-$low) or only somewhat ($\delta-$medium) beneficial, while there is no effect among those who expect the treatment to have a strong positive effect ($\delta-$high, see Panel C). Since only very few individuals expect the treatment to have a negative impact on their labor market outcomes, those participants who actually have weaker incentives adjust the expected participation probability π based on the expected treatment effect δ are responsible for the effect of π on the employment probability. Moreover, as shown in Panel D, conditioning on the expected treatment effect δ has only a minor impact on the estimated average effect of the expected participation probability. It can be concluded that the endogenous adjustment of the expected participation probability according to the expected treatment effect is unlikely to explain the baseline results.

[Insert Table 9 about here]

5 Conclusion

The paper analyzes the role of subjective expectations of unemployed workers about the future participation in labor market training for the program effectiveness. Based on a unique combination of German survey and administrative data, I can show that expecting the treatment ex ante increases the effectiveness of human capital-intensive long-term training programs. Since the program assignment is organized through a voucher system and participants can freely choose a program provider, those individuals who expect to actually participate have strong incentives to search for an appropriate program in order to increase the match quality and therefore also effectiveness of the treatment. For participants in short-term training, who cannot choose the program provider, the pre-treatment expectations are unrelated the effectiveness of the intervention. A variety of balancing tests and sensitivity analyses support the causal interpretation of the results, while participants who expected the treatment also state ex post that they invested greater effort related to the program preparation and that they are more satisfied with the treatment.

Moreover, I use a causal forest algorithm to identify background characteristics that are particularly relevant for the effect heterogeneity. The findings show that the individual's job search effort, the degree to which people believe that they have control over life outcomes

(locus of control) and their labor market histories are particularly important predictors for the effect of the expected participation probability on the program effectiveness. The results are in line with theoretical considerations as the effect is the strongest for those groups who can be assumed to have the strongest incentives to change their behavior based on their expectations, respectively the highest returns with respect to the behavioral adjustment. The findings have important policy implications as they provide first evidence that the effectiveness of labor market programs relying on market-based assignment procedures can depend on the participants' expectations about a treatment. Therefore, policy makers who aim to utilize mechanisms that involve high degrees of competition and individual choices should note that the counseling of unemployed workers and provision of all relevant information has great importance for the effectiveness of an intervention. Given that many countries nowadays rely on market-based mechanisms for various kinds of policies, the relationship between the underlying incentive structures, subjective perceptions and optimal information strategies is surely a worthwhile focus for future research.

References

Altmann, S., A. Falk, S. Jäger, and F. Zimmermann (2017): "Learning about job search: A field experiment with job seekers in Germany," *forthcoming: Journal of Public Economics*.

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

Athey, S. and G. Imbens (2016): "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, 113, 7353–7360.

Athey, S. and S. Wager (2019): "Estimating Treatment Effects with Causal Forests: An Application," Working Paper, Stanford University, *arXiv preprint arXiv:1902.07409*.

Behaghel, L., B. Crépon, and M. Gurgand (2014): "Private and public provision of counseling to job seekers: Evidence from a large controlled experiment," *American Economic Journal: Applied Economics*, 6, 142–174.

Behncke, S., M. Frölich, and M. Lechner (2010a): "A caseworker like me–does the similarity between the unemployed and their caseworkers increase job placements?" *The Economic Journal*, 120, 1430–1459.

——— (2010b): "Unemployed and their caseworkers: should they be friends or foes?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173, 67–92.

Bernhard, S. and T. Kruppe (2012): "Effectiveness of further vocational training in Germany: Empirical findings for persons receiving means-tested unemployment benefits," *Schmollers Jahrbuch*, 132, 501–526.

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, 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*, 94, 1313–1327.

Bottazzi, R., T. Jappelli, and M. Padula (2006): "Retirement expectations, pension reforms, and their impact on private wealth accumulation," *Journal of Public Economics*, 90, 2187–2212.

Breiman, L. (2001): "Random forests," *Machine learning*, 45, 5–32.

Caliendo, M., D. Cobb-Clark, and A. Uhlendorff (2015): "Locus of control and job search strategies," *Review of Economics and Statistics*, 97, 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., R. Mahlstedt, and O. A. Mitnik (2017): "Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies," *Labour Economics*, 46, 14–25.

Card, D., J. Kluve, and A. Weber (2010): "Active labour market policy evaluations: A meta-analysis," *The Economic Journal*, 120, F452–F477.

——— (2017): "What works? A meta-analysis of recent active labor market program evaluations," *forthcoming: Journal of the European Economic Association.*

Carling, K. and L. Larsson (2005): "Does early intervention help the unemployed youth?" *Labour Economics*, 3, 301–319.

Chan, S. and A. H. Stevens (2004): "Do changes in pension incentives affect retirement? A longitudinal study of subjective retirement expectations," *Journal of Public Economics*, 88, 1307–1333.

Chetty, R. and E. Saez (2013): "Teaching the tax code: Earnings responses to an experiment with EITC recipients," *American Economic Journal: Applied Economics*, 5, 1–31.

Christelis, D., D. Georgarakos, T. Jappelli, and M. van Rooij (2016): "Trust in the central bank and inflation expectations," Working paper.

Crépon, B., M. Ferracci, G. Jolivet, and G. J. van den Berg (2018): "Information shocks and the empirical evaluation of training programs during unemployment spells," *Journal of Applied Econometrics*, 33, 594–616.

D'Acunto, F., D. Hoang, and M. Weber (2015): "Inflation Expectations and Consumption Expenditure," Working paper.

DellaVigna, S., A. Lindner, B. Reizer, and J. Schmieder (2017): "Referenced-dependent job search: Evidence from Hungary," *The Quarterly Journal of Economics*, 13, 1969–2018.

DellaVigna, S. and D. Paserman (2005): "Job search and impatience," *Journal of Labor Economics*, 23, 527–588.

Doerr, A., B. Fitzenberger, T. Kruppe, M. Paul, and A. Strittmatter (2017): "Employment and earnings effects of awarding training vouchers in Germany," *Industrial and Labor Relations Review*, 70, 767–812.

Eberle, J., R. Mahlstedt, and A. Schmucker (2017): "IZA/IAB Linked Evaluation Dataset 1993-2010," FDZ Data Report 02/2017.

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*, 91/92, 321–355.

Gatz, M. and M. J. Karel (1993): "Individual change in perceived control over 20 years," *International Journal of Behavioral Development*, 16, 305–322.

Geerdsen, L. (2006): "Is there a threat effect of labour market programmes? A study of ALMP in the Danish UI system," *The Economic Journal*, 116, 738–750.

Geerdsen, L. and A. Holm (2007): "Duration of UI periods and the perceived threat effect from labour market programmes," *Labour Economics*, 14, 639–652.

Gorter, C. and G. Kalb (1996): "Estimating the effect of counseling and monitoring the unemployed using a job search model," *Journal of Human Resources*, 31, 590–610.

Graversen, B. and J. Van Ours (2008): "How to help unemployed find jobs quickly: Experimental evidence from a mandatory activation program," *Journal of Public Economics*, 92, 2020–2035.

——— (2011): "An activation program as a stick to job finding," *Labour*, 25, 167–181.

Hägglund, P. (2011): "Are there pre-programme effects of active placement efforts? Evidence from a social experiment," *Economics Letters*, 112, 91–93.

Heineck, G. and S. Anger (2010): "The returns to cognitive abilities and personality traits in Germany," *Labour Economics*, 17, 535–546.

Hipp, L. and M. E. Warner (2008): "Market forces for the unemployed? Training vouchers in Germany and the USA," *Social Policy & Administration*, 42, 77–101.

Jacobi, L. and J. Kluve (2007): "Before and after the Hartz reforms: The performance of active labour market policy in Germany," *Journal for Labour Market Research*, 40, 45–64.

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

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, 742–784.

Lechner, M. and J. Smith (2007): "What is the value added by caseworkers?" *Labour economics*, 14, 135–151.

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, 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, 184–211.

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

Osikominu, A. (2012): "Quick job entry or long-term human capital development? The dynamic effects of alternative training schemes," *Review of Economic Studies*, 80, 313–342.

Paserman, D. (2008): "Job search and hyperbolic discounting: Structural estimation and policy evaluation," *The Economic Journal*, 118, 1418–1452.

Rinne, U., A. Uhlendorff, and Z. Zhao (2013): "Vouchers and caseworkers in training programs for the unemployed," *Empirical Economics*, 45, 1089–1127.

Rosholm, M. and M. Svarer (2008): "The threat effect of active labour market programmes," *The Scandinavian Journal of Economics*, 110, 385–401.

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

Schütz, H., P. Kupka, S. Koch, and B. Kaltenborn (2011): "Eingliederungsvereinbarungen in der Praxis: Reformziele noch nicht erreicht," Iab-kurzbericht 18/2011.

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

Sol, E., M. Westerveld, and M. Westerveld (2005): *Contractualism in employment services: a new form of welfare state governance*, vol. 29, Kluwer Law International BV.

Spinnewijn, J. (2015): "Unemployed but optimistic: Optimal insurance design with biased beliefs," *Journal of the European Economic Association*, 13, 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, 606–616.

van den Berg, G., A. Bozio, and M. Costa Dias (2014a): "Policy discontinuity and duration outcomes," IZA Discussion Paper 8450.

van den Berg, G., B. Hofmann, G. Stephan, and A. Uhlendorff (2014b): "Was Vermittlungsfachkräfte von Eingliederungsvereinbarungen halten: Befragungsergebnisse aus einem Modellprojekt," IAB-Forschungsbericht 11/2014.

Wager, S. and S. Athey (2018): "Estimation and inference of heterogeneous treatment effects using random forests," *Journal of the American Statistical Association*, 113, 1228–1242.

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

Tables and Figures

Table 1: Descriptive Statistics on Labor Market Outcomes

	Expected participation probability		
	$\pi-$low	$\pi-$high	$P-$value
A. Long-term training			
No. of observations	206	501	
Regular employed in month			
$t+12$	0.25	0.38	0.00
$t+30$	0.50	0.63	0.00
Average monthly earnings in €	1,246	1,093	0.32
B. Short-term training			
No. of observations	486	971	
Regular employed in month			
$t+12$	0.50	0.44	0.04
$t+30$	0.59	0.59	0.98
Average monthly earnings in €	1,104	1,054	0.45
C. Non-participants			
No. of observations	2,163	2,138	
Regular employed in month			
$t+12$	0.53	0.55	0.07
$t+30$	0.57	0.59	0.27
Average monthly earnings in €	1,014	987	0.36

Note: Percentage share unless indicated otherwise. $P-$values measured based on two-tailed t-tests on equal means. Outcome variables are measured relative to the month of entry into unemployment t.

Table 2: Marginal Effects of Probit Estimation: Selection into Realized Treatment

	Long-term training v. non-participation	Short-term training v. non-participation	Long-term training v. short-term training
	(1)	(2)	(3)
Expected part. probability: high ($\pi \geq 5$)	0.0826***	0.0865***	0.0257
	(0.0125)	(0.0129)	(0.0216)
No. of observations	5,008	5,758	2,164
log-Likelihood	-1886.6	-3137.4	-1266.4
$P-$value (LR-test)	0.0000	0.0000	0.0000
Pseudo-R^2	0.0746	0.0367	0.0738
Control variables			
Socio-demographic characteristics	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes
Labor market histories	Yes	Yes	Yes
Regional and seasonal information	Yes	Yes	Yes
Personality traits	Yes	Yes	Yes

Note: Depicted are average marginal effects based on probit models estimating the impact of the expected on the realized participation probability. Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level. Full estimation results are available in Table B.2 in the Appendix.

Table 3: Difference-in-Difference Model: Balancing Test wrt Observed Covariates

	Effect of expected assignment probability π-high v. π-low			
	Unconditional		IPW	
	Coef.	SE	Coef.	SE
Female	0.0928*	(0.0498)	0.0535	(0.0572)
Age				
16-24 years	-0.0092	(0.0414)	-0.0391	(0.0417)
25-34 years	-0.0295	(0.0436)	0.0053	(0.0484)
35-44 years	0.1084**	(0.0446)	0.0180	(0.0561)
45-55 years	-0.0697*	0.0422	0.0158	0.0480
School leaving degree (Ref.: None)				
Lower sec. degree	-0.0483	(0.0461)	-0.0268	(0.0486)
Middle sec. degree	0.0568	(0.0496)	0.0214	(0.0580)
Upper sec. degree	0.0022	(0.0405)	0.0090	(0.0484)
Higher education (Ref.: None)				
Internal/external prof. training	-0.0414	(0.0428)	-0.0303	(0.0492)
University degree	0.0398	(0.0360)	0.0355	(0.0424)
German citizenship	-0.0061	(0.0225)	0.0422	(0.0431)
Migration background	-0.0243	(0.0352)	-0.0543	(0.0507)
Married (or cohabiting)	0.1028**	(0.0488)	0.0398	(0.0574)
Problems with childcare	0.0632	(0.0970)	-0.0342	(0.1360)
Partner is full-time employed	0.0011	(0.0211)	-0.0053	(0.0246)
Children in household				
Age 0-3 years	0.0038	(0.0307)	0.0057	(0.0330)
Age 4-6 years	0.0273	(0.0263)	0.0298	(0.0290)
Age 7-15 years	0.0605	(0.0381)	0.0249	(0.0464)
Age 16-18 years	-0.0027	(0.0279)	0.0062	(0.0299)
Searching for full-time employment	0.0943***	(0.0339)	0.0488	(0.0358)
Region				
West-Germany & UE rate 0-6%	0.0707	(0.0471)	-0.0380	(0.0575)
West-Germany & UE rate 6+%	-0.0350	(0.0478)	-0.0136	(0.0545)
East-Germany & UE rate 9-14%	0.0231	(0.0368)	0.0477	(0.0344)
East-Germany & UE rate 15+%	-0.0588*	(0.0346)	0.0038	(0.0362)
Entry into unemployment (Ref.: 2nd quarter 2007)				
3rd quarter 2007	0.0363	(0.0396)	-0.0057	(0.0471)
4th quarter 2007	-0.0091	(0.0449)	0.0381	(0.0476)
1st quarter 2008	-0.0583	(0.0444)	-0.0688	(0.0547)
2nd quarter 2008	0.0232	(0.0370)	0.0161	(0.0424)
Time to interview				
7 weeks	-0.0094	(0.0122)	0.0012	(0.0101)
8 weeks	-0.0253	(0.0432)	-0.0316	(0.0493)
9 weeks	0.0018	(0.0419)	0.0233	(0.0437)
10 weeks	-0.0248	(0.0377)	-0.0085	(0.0496)
11 weeks	0.0361	(0.0367)	0.0121	(0.0416)
12 weeks	-0.0214	(0.0271)	-0.0059	(0.0295)
13 weeks	0.0325*	(0.0176)	0.0070	(0.0259)
14 weeks	0.0105	(0.0235)	0.0024	(0.0232)
Unemployment benefit recipient	(0.0349)	0.0381	-0.0087	(0.0411)
Last daily income in €	-1.2772	(3.1967)	0.5645	(3.5257)
Employment status before unemployment (Ref.: Other)				
Regular employment	-0.0769*	(0.0440)	-0.0082	(0.0487)
Subsidized employment	0.0139	(0.0394)	-0.0076	(0.0425)
Last job was full-time employment	-0.0200	(0.0431)	0.0125	(0.0461)
Months in employment				
in last year	0.4743	(0.4612)	0.0690	(0.5090)
in last 5 years	2.0820	(1.8653)	-0.6831	(2.1805)
in last 10 years	3.8462	(3.5512)	1.3529	(4.1941)
Months in unemployment				
in last year	-0.1893	(0.1725)	-0.0746	(0.1862)
in last 5 years	-1.4042*	(0.7452)	0.3888	(0.8685)
in last 10 years	-1.7865	(1.5033)	0.9238	(1.5544)
Openness	0.0038	(0.0995)	-0.0266	(0.1143)
Conscientiousness	0.0925	(0.0932)	-0.0091	(0.0943)
Extraversion	0.1178	(0.0988)	-0.0271	(0.1098)
Neuroticism	0.0200	(0.0973)	0.0554	(0.1120)
Locus of Control	0.0650**	(0.0291)	0.0145	(0.0392)

Note: Depicted are differences-in-differences (DID) between participants in long-term training and short-term training with high, respectively low expected assignment probabilities π with respect to observed characteristics: 1) unconditional DID and 2) DID based on inverse probability weighting (IPW). Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level.

Table 4: Balancing Test with Respect to Job Search and Expected Labor Market Outcomes

	Effect of expected participation probability π-high v. π-low		
Pre-treatment outcome	Long-term training (1)	Short-term training (2)	Difference-in-Difference (3)
A. Job search behavior			
Log average weekly no. of applications	-0.0533	-0.0129	-0.0404
	(0.0570)	(0.0324)	(0.0656)
No. of search channels	0.0647	0.0882	-0.0235
	(0.1787)	(0.0940)	(0.2019)
B. Expected labor market outcomes			
Expected reemployment probability within six months			
very low	-0.0409**	-0.0274**	-0.0135
	(0.0178)	(0.0111)	(0.0209)
low	-0.0334	-0.0593***	0.0259
	(0.0284)	(0.0178)	(0.0335)
high	0.0834*	0.0891***	-0.0057
	(0.0451)	(0.0288)	(0.0535)
very high	-0.0000	-0.0089	0.0089
	(0.0486)	(0.0297)	(0.0570)
Log expected monthly net income	-0.0286	-0.0206	-0.0080
	(0.0346)	(0.0228)	(0.0414)
Assumed influence of employment agency on job finding			
will improve reemployment prospects	0.1665***	0.1503***	0.0163
	(0.0418)	(0.0289)	(0.0508)
will worsen reemployment prospects	-0.0254	-0.0253*	-0.0001
	(0.0181)	(0.0148)	(0.0234)
No. of observations	707	1,457	2,164
Control variables			
Socio-demographic characteristics	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes
Labor market histories	Yes	Yes	Yes
Regional and seasonal information	Yes	Yes	Yes
Personality traits	Yes	Yes	Yes

Note: Depicted are average treatment effects of the expected assignment probability π-high (relative to π-low) based on inverse probability weighting (IPW). Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level.

Table 5: Baseline Results: Expected Participation Probability and Realized Labor Market Outcomes

	Effect of expected participation probability π-high v. π-low					
	Long-term training	Short-term training	Non-participants	Difference-in-Difference$^{(a)}$		
	IPW (1)	IPW (2)	IPW (3)	DID (4)	DDD_{t-2} (5)	DDD_{t-5} (6)
Regular employed in month						
$t+12$	0.1211***	-0.0682**	0.0239	0.1893***	0.2210***	0.1997***
	(0.0413)	(0.0300)	(0.0158)	(0.0510)	(0.0651)	(0.0635)
$t+30$	0.1033**	0.0005	0.0169	0.1028*	0.1345**	0.1131*
	(0.0476)	(0.0299)	(0.0157)	(0.0562)	(0.0611)	(0.0584)
Average monthly earnings in €	-29.27	-54.26	-22.11	24.99	19.54	
	(158.69)	(66.67)	(31.73)	(172.06)	(172.66)	
No. of observations	707	1,457	4,301	2,164	2,164	2,164
Control variables						
Socio-demographic characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Labor market histories	Yes	Yes	Yes	Yes	Yes	Yes
Regional and seasonal information	Yes	Yes	Yes	Yes	Yes	Yes
Personality traits	Yes	Yes	Yes	Yes	Yes	Yes

Note: Depicted are average treatment effects of the expected participation probability π-high (relative to π-low) based on inverse probability weighting (IPW). Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level.
$^{(a)}$The difference-in-difference model (column 4-6) shows the impact of π-high on participants in long-term training relative to the impact of π-high on participants in short-term training. DDD_{t-2}/DDD_{t-5} refers to the conditional difference-in-difference-in-difference model, where the reference level is given by the average employment rate within the last two/five years before the entry into unemployment. For average earnings the reference level of the conditional difference-in-difference-in-difference model is given by the last income before becoming unemployed.

Table 6: Descriptive Statistics on Assessment of Long-term Training Programs

	Expected participation probability		
	$\pi-$low	$\pi-$high	$P-$value
No. of observations	133	301	
A. Effort related to program preparation			
Initiative to participate stem from			
unemployed	0.62	0.58	0.38
caseworker	0.14	0.09	0.16
both sides	0.23	0.33	0.05
Participant had influence on			
type of program	0.38	0.48	0.05
timing of program	0.22	0.31	0.06
Placement through			
Employment agency	0.65	0.66	0.77
Other institution	0.10	0.04	0.03
Social network	0.08	0.14	0.06
B. Assessment of program quality			
Employment opportunities due to participation			
strongly improved	0.23	0.33	0.04
somewhat improved	0.49	0.38	0.03
not influenced	0.26	0.27	0.95
worsened	0.01	0.02	0.46
Satisfaction with program$^{(a)}$	6.82	7.32	0.07
C. Completion of program			
Participant received degree	0.80	0.79	0.74

Note: Descriptive statistics with respect to ex post assessment of program participation for subsample of participants interviewed at second survey wave about one year after entry into unemployment. Percentage share unless indicated otherwise. $P-$values measured based on two-tailed t-tests on equal means.
$^{(a)}$Measured on a scale from 0 ($\equiv$ not satisfied at all) to 10 ($\equiv$ fully satisfied).

Table 7: Subjective Expectations and Program Effectiveness by Motivation to Contact Employment Agency

	Effect of expected participation probability π-high v. π-low	
	Long-term training	
	Motivation to contact agency	
	Low	High
	IPW	IPW
	(1)	(2)
Regular employed in month		
$t+12$	0.0954*	0.1578***
	(0.0570)	(0.0599)
$t+30$	0.0580	0.1699**
	(0.0630)	(0.0691)
Average monthly earnings in €	-177.49	215.52**
	(259.90)	(108.08)
No. of observations	379	328
Control variables		
Socio-demographic characteristics	Yes	Yes
Household characteristics	Yes	Yes
Labor market histories	Yes	Yes
Regional and seasonal information	Yes	Yes
Personality traits	Yes	Yes

Note: Depicted are average treatment effects of the expected participation probability π-high (relative to π-low) based on inverse probability weighting (IPW) separated for individuals with a low (below median) and high (above median) predicted probability to contact the employment agency to participate in a training program. Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level.

Table 8: Comparing Predicted and Realized Effects of the Expected Participation Probability

	Effect of expected part. probability π-high v. π-low		
	Within-sample	**Out-of-sample**	
	Long-term training	Short-term training	Non-participants
	IPW	IPW	IPW
Outcome variable: Regular employed in month $t+30$			
A. Low predicted effect: $\hat{\tau}_i^{CF} \leq \hat{\tau}_{0.5}^{CF}$	-0.2133***	0.0279	0.0085
	(0.0564)	(0.0415)	(0.0220)
No. of observations	354	729	2,151
Mean π-low	0.6847	0.5885	0.5913
B. High predicted effect: $\hat{\tau}_i^{CF} > \hat{\tau}_{0.5}^{CF}$	0.4438***	-0.0264	0.0249
	(0.0595)	(0.0432)	(0.0223)
No. of observations	353	728	2,150
Mean π-low	0.2737	0.5885	0.5459
Control variables			
Socio-demographic characteristics	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes
Labor market histories	Yes	Yes	Yes
Regional and seasonal information	Yes	Yes	Yes
Personality traits	Yes	Yes	Yes

Note: Depicted are average treatment effects of the expected assignment probability π-high (relative to π-low) based on inverse probability weighting (IPW) for subgroups with different levels of expected treatment effects. Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level.

[a] Panel D shows the weighted average of the separate estimates for the three sub-samples from Panel A-C with weights calculated based on the average number of observation within each sample. Standard errors are obtained based on bootstrapping with 999 replications.

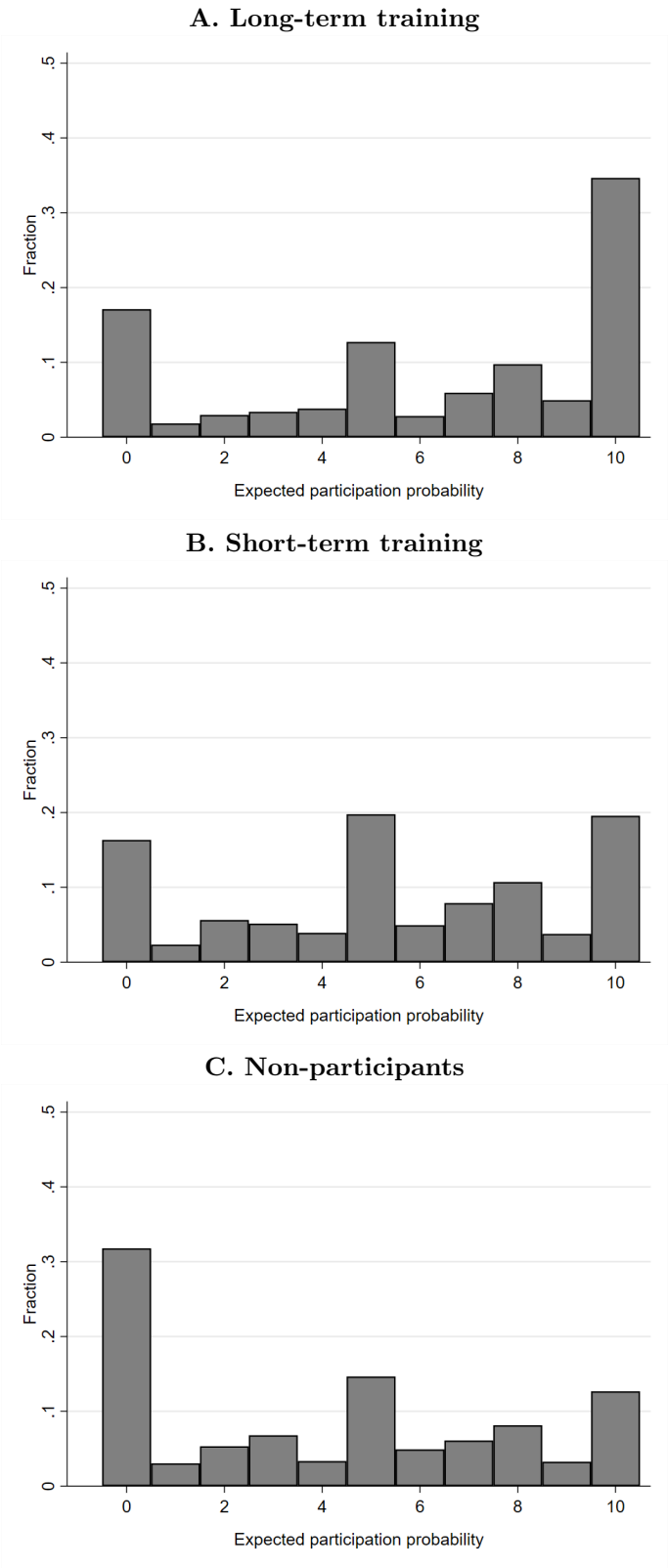
Table 9: Interdependence of Expected Participation Probability and Expected Treatment Effect

	Effect of expected participation probability π-high v. π-low	
	Long-term training	Short-term training
	IPW (1)	IPW (2)
Outcome variable: Regular employed in month $t+30$		
A. Expected treatment effect δ: low	0.1690*	0.0640
	(0.1016)	(0.0562)
No. of observations	116	370
Mean π-low	0.4783	0.5723
B. Expected treatment effect δ: medium	0.1820**	-0.0388
	(0.0712)	(0.0453)
No. of observations	266	633
Mean π-low	0.4634	0.6345
C. Expected treatment effect δ: high	-0.0098	-0.0009
	(0.0890)	(0.0569)
No. of observations	325	454
Mean π-low	0.5636	0.5366
D. Full sample conditioned on $\delta^{(a)}$	0.0917*	-0.0009
	(0.0510)	(0.0308)
No. of observations	707	1,457
Mean π-low	0.4951	0.5885
Control variables		
Socio-demographic characteristics	Yes	Yes
Household characteristics	Yes	Yes
Labor market histories	Yes	Yes
Regional and seasonal information	Yes	Yes
Personality traits	Yes	Yes

Note: Depicted are average treatment effects of the expected assignment probability π-high (relative to π-low) based on inverse probability weighting (IPW) for subgroups with different levels of expected treatment effects. Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level.

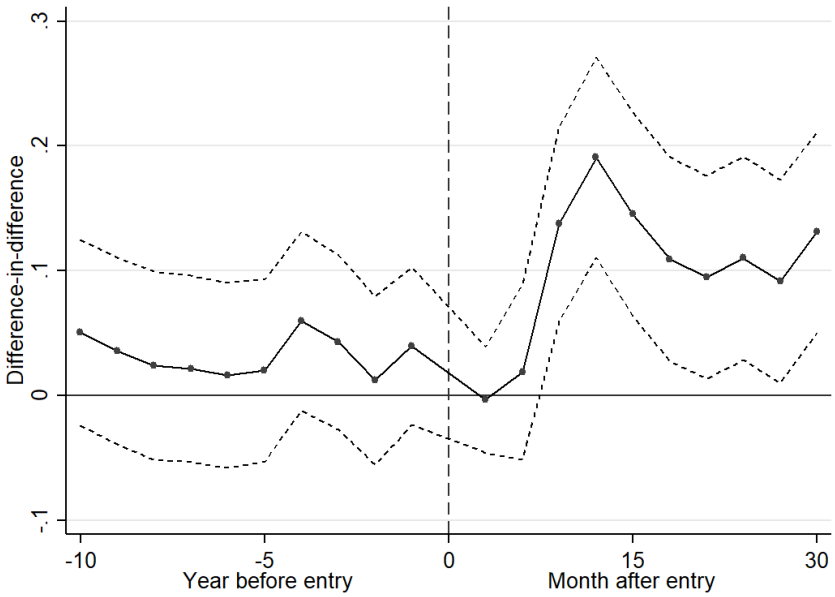
$^{(a)}$Panel D shows the weighted average of the separate estimates for the three sub-samples from Panel A-C with weights calculated based on the average number of observation within each sample. Standard errors are obtained based on bootstrapping with 999 replications.

Figure 1: Distribution of the Expected Participation Probability



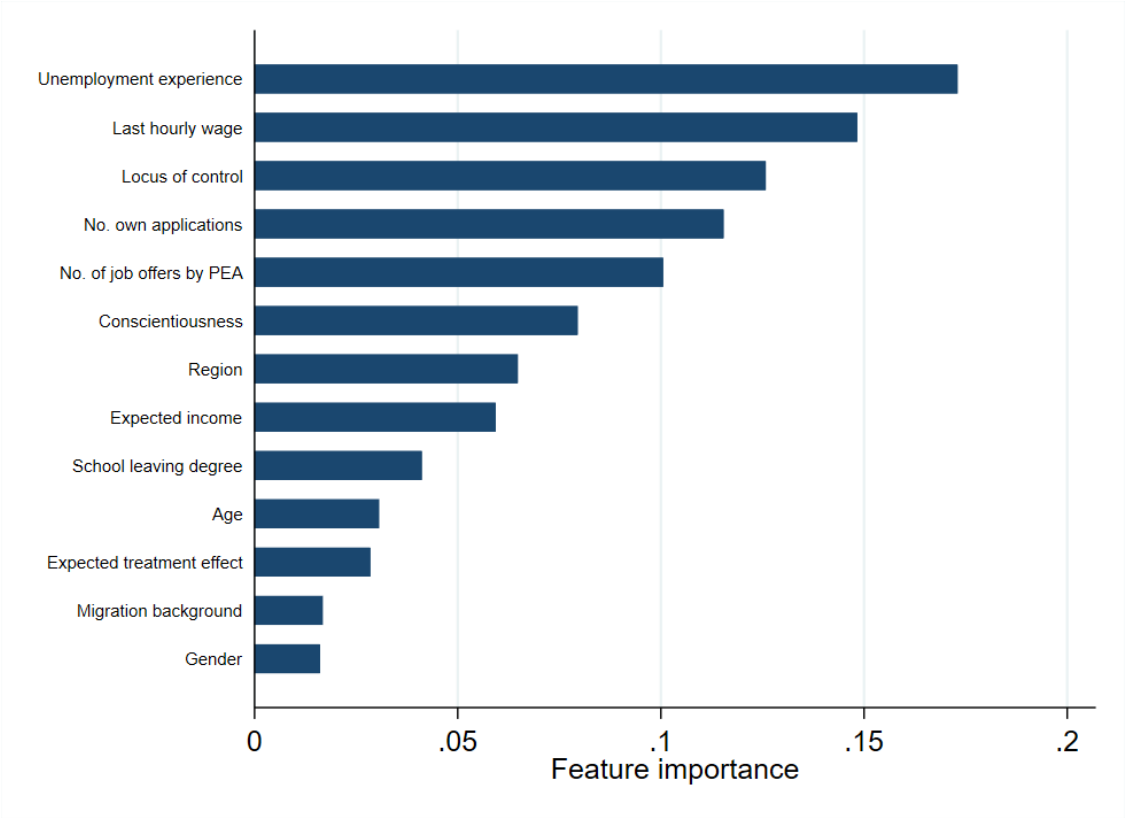
Depicted are answers to the question: *"Assuming that you are still unemployed during the next 3 months. What is the probability that you will participate in a training scheme?"* 0 = very unlikely; 10 = very likely.

Figure 2: Unconditional Difference-in-Differences Over Time



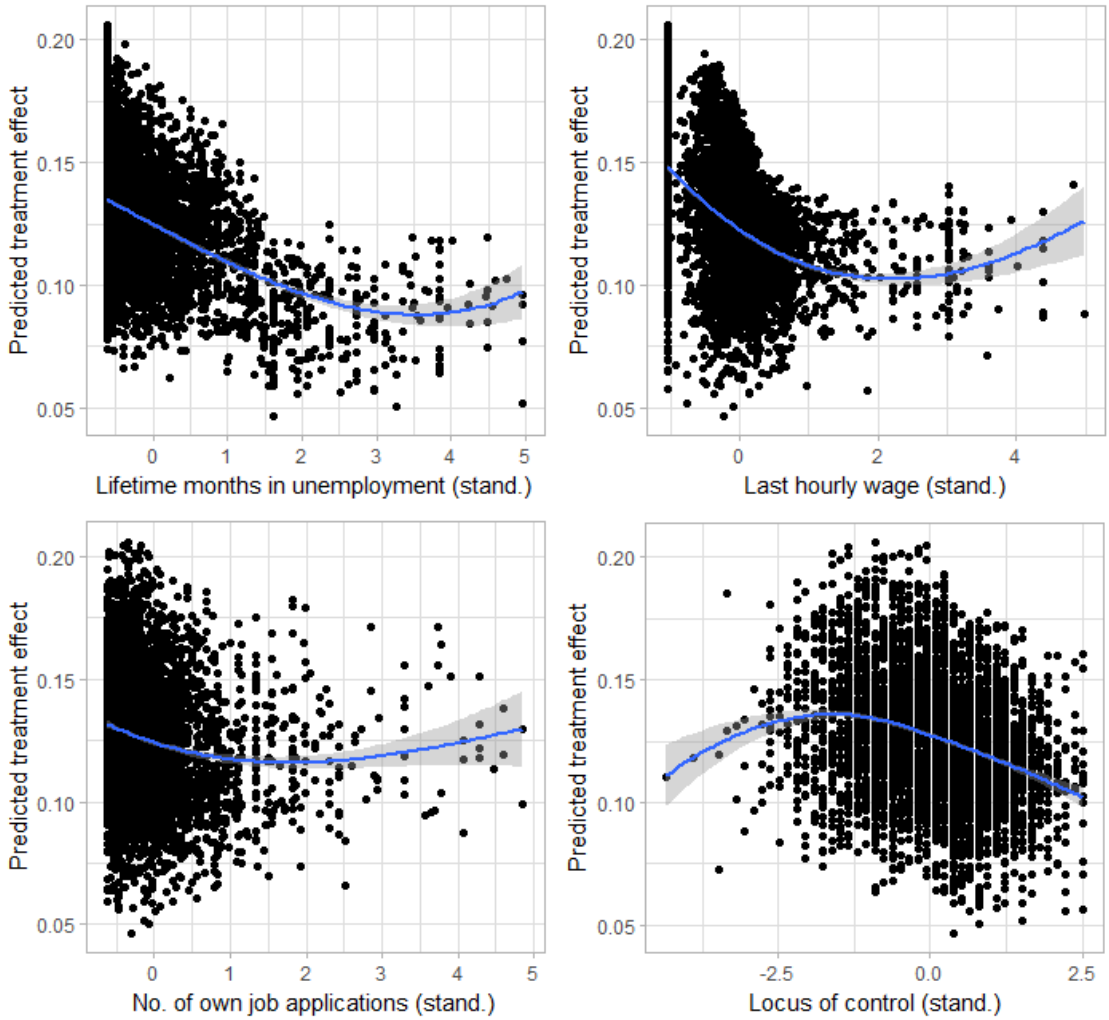
Note: Depicted are unconditional differences-in-differences (DID) referring to the impact of π-high on participants in long-term training relative to the impact of π-high on participants in short-term training over time (solid line) and the corresponding 90% confidence interval (dashed line).
The left-hand side shows differences-in-differences in average yearly employment rates for the last 10 years before the entry into unemployment. The right-hand side shows differences-in-differences in monthly employment rates for a period of 30 months after the entry into unemployment.

Figure 3: Feature Importance for Final Specification



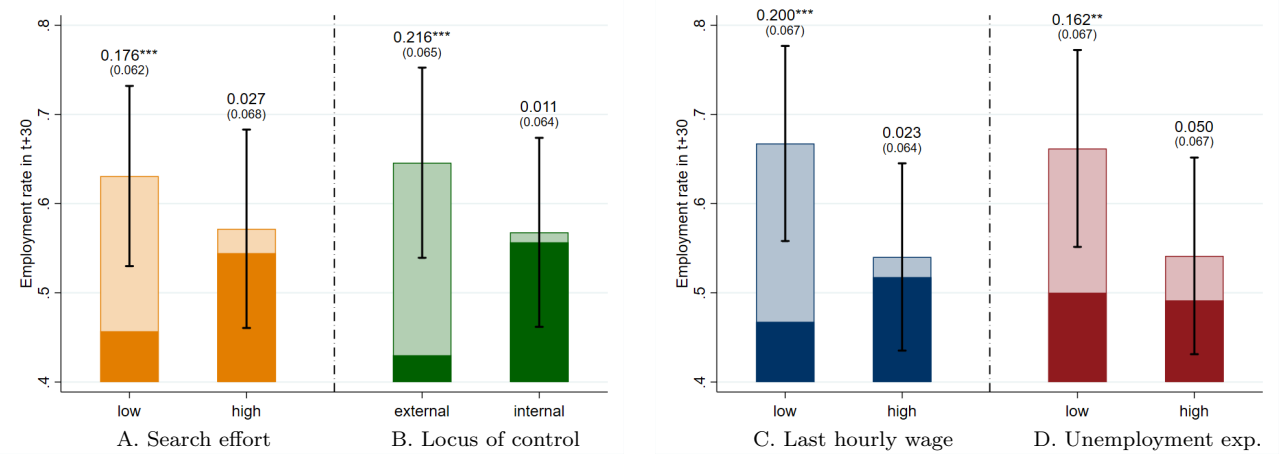
Note: Feature importance is defined as the normalized average reduction in the expected mean squared error after including a specific variable. Variables are listed in descending order of feature importance. Variables included in the final specification are pre-selected from the full model (see Figure B.3).

Figure 4: Conditional Average Treatment Effects by Most Important Features



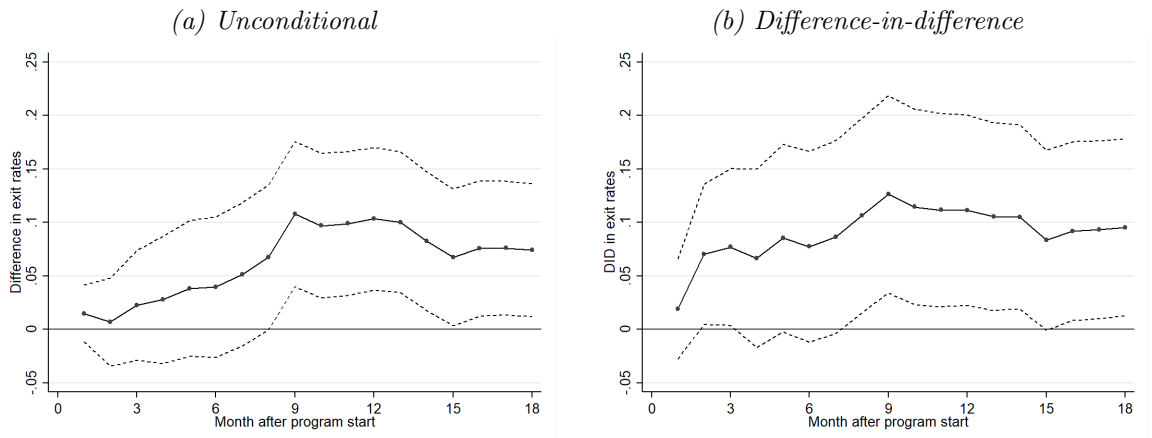
Note: Depicted are conditional average treatment effects for the four most important features identified based on the causal forest approach. In total, the model includes 13 features as depicted in Figure B.3.

Figure 5: Average Treatment Effects: Effect Heterogeneity by Most Important Features



Note: The numbers on top of the bars depict the average treatment effects of the expected participation probability π-high (relative to π-low) with standard errors in parenthesis for separated estimation samples divided at the median of the four most important features derived from the causal forest approach. The outcome variable is regular employment 30 months after the entry into unemployment.
The intense colored region denotes mean employment rates for individuals reporting low expected participation probabilities (π-low), while the light colored region depicts the average treatment effects of π-high including the corresponding 90% confidence intervals.

Figure 6: Impact of the Expected Participation Probability π Relative to Program Start



Note: Depicted are differences in cumulated monthly exit rates from unemployment to regular employment relative to beginning of the program.
(a) Unconditional differences between participants in long-term training with high expected assignment probabilities (π-high) and low expected assignment probabilities (π-low).
(b) Differences-in-differences between participants in long-term training and short-term training with high, respectively low expected assignment probabilities π based on inverse probability weighting (IPW). *Control variables:* socio-demographic characteristics, household characteristics, labor market histories, regional/seasonal information and personality traits.

Online Appendix

Subjective Expectations and the Effectiveness of Labor Market Policies

Robert Mahlstedt*

July 14, 2019

Section A shows a job search model accounting for subjective expectations about the enrollment in a labor market program.

Section B provides additional tables and figures supplementing the descriptive statistics in Section 2.5, the discussion in Section 3.2 and the estimation results in Section 4.

*University of Copenhagen, Denmark; IZA Bonn, Germany; robert.mahlstedt@econ.ku.dk
Corresponding address: University of Copenhagen, Department of Economics, Øster Farimagsgade 5, 1353 København K, Denmark. Tel: +45 35 33 79 22

A Theoretical Framework

To understand the role of subjective expectations about a future program participation and show the potential mechanisms through which they affect the individual behavior of unemployed workers, it is useful to consider a search model where job seekers face the possibility to participate in a training program in the future (see also van den Berg *et al.*, 2009). Consider an unemployed individual who searches sequentially for a new job. In each period t, the job seeker decides about the search strategy s_t, which determines the probability to find a new job $\lambda(s_t)$. Finding a job implies utility ω, which is assumed to always exceed the value of remaining unemployed. Moreover, it is assumed that the job seeker takes into account the possibility of participating in a training program during the unemployment spell. Given that she remains unemployed, she expects to be enrolled in the program in period $t+1$ with probability π. However, in contrast to the baseline model by van den Berg *et al.* (2009), the agent has to decide about a second dimension of effort z_t, capturing all activities that will prepare her for the treatment. For instance, this could involve the search for an appropriate program or program provider, the acquisition of relevant materials or adaptations of the job search strategy that might be necessary due to the treatment. Both types of effort are assumed to generate costs $c(s_t, z_t)$. Hence, for a given discount rate ρ, the inter-temporal value of being unemployed is given as:

$$V_t^u = -c(s_t, z_t) + \rho \left\{ \lambda(s_t)\omega + (1 - \lambda(s_t)) \left((1-\pi)V_{t+1}^u + \pi V_{t+1}^p \right) \right\}, \tag{A.1}$$

where V_{t+1}^p denotes the inter-temporal value of being enrolled in a program in the next period. After the start of the program, the agent no longer faces the possibility of a future treatment, but the effort that she has devoted into the program preparation positively affects the job finding prospects after the program start.[23] Therefore, the inter-temporal value of being treated is given as:

$$V_t^p = -k(s_t) + \rho \left\{ \mu(s_t, z_{t-1})\omega + (1 - \mu(s_t, z_{t-1}))V_{t+1}^p \right\}, \tag{A.2}$$

where $k(s_t)$ denotes the search costs when being enrolled, which might be different than during unemployment, while μ characterizes the job finding prospect after the program

[23] For the ease of notation, is assumed that the quality of the provider has positive effect on the fob finding prospects from the first day of the treatment and does not require the successful completion of the treatment.

start, which can be interpreted as the effectiveness of the treatment and depends on the current search effort and the previous effort spent into the program preparation.

The optimal behavior of an individual who is unemployed and has not yet been enrolled is characterized by the following conditions, which imply that the agent equalizes the marginal costs and returns with respect to both types of effort:

$$c_s'(s_t, z_t) = \rho\lambda_s'(s_t)\left\{\omega - (1-\pi)V_{t+1}^u - \pi V_{t+1}^p\right\} \tag{A.3}$$

$$c_z'(s_t, z_t) = \rho^2\pi(1-\lambda(s_t))\mu_z'(s_{t+1}, z_t)\left\{\omega - V_{t+1}^p\right\} \tag{A.4}$$

The left-hand side denotes marginal changes of the search costs with respect to the two different types of effort, while the right-side of each condition characterizes the marginal return with respect to job search (equation A.3), respectively preparation effort (equation A.4). Therefore, the impact of the expected participation probability π on the individual effort allocation has the following properties:

$$\frac{\partial z_t}{\partial \pi} > 0 \quad \text{and} \quad \frac{\partial s_t}{\partial \pi} < 0$$

if the following condition holds:

$$\mu_z'(s_{t+1}, z_t) > \frac{\lambda_s'(s_t)}{1-\lambda(s_t)} \cdot \frac{V_{t+1}^u - V_{t+1}^p}{\omega - V_{t+1}^p}. \tag{A.5}$$

It implies that the agent will raise the effort spent into program preparation to make the treatment more effective if the (assumed) marginal increase of the program effectiveness μ_z' is sufficiently large. Since an increase of the preparation effort is associated with reduction of the job search effort (as time and resources are limited) and only pays off if the agent remains unemployed and actually enters the program, the likelihood that condition A.5 holds decreases with an increasing total job finding rate $\lambda(s_t)$ and the marginal impact of the search effort $\lambda_s'(s_t)$. If this is not the case, the agent has incentives to raise her job search effort as a reaction to an increasing participation probability in order to leave unemployment and avoid the treatment. Therefore, the impact of the expected participation probability π on the program effectiveness crucially depends on the marginal impact of the preparation effort.

B Supplementary Tables and Figures

Table B.1 shows the distribution of expected treatment effects by the expected participation probability.

Table B.2 shows marginal effects of probit models estimating actual participation probability conditioned on the subjective expectations and the full set of covariates. The first, respectively second specification estimates the probability to participate in long-term, respectively short-term training, relative to non-participation. The third specification directly compares the likelihood of participating in one of the two programs.

Table B.3 shows the results of a sensitivity analysis using an alternative expectation measure. In this specification π refers to answer on the survey question: *"Assuming that you are still unemployed during the next three months. What do you think is the probability that you will participate in <u>any scheme that aims to improve you chances for employment</u>?"*

Figure B.1 illustrates the underlying economic framework, empirical setting and the timing of the measurement of the relevant variables.

Figure B.2 shows the distribution of the expected treatment effects δ.

Figure B.3 shows the feature importance for the full set of covariates included in the pilot causal forest.

Figure B.4 shows the distribution of the predicted conditional average treatment effects based on the causal forest algorithm.

Table B.1: Expected Treatment Effect by Expected Participation Probability and Realized Treatment Status

	Expected participation probability		
	$\pi-$low	$\pi-$high	$P-$value
A. Long-term training			
No. of observations	206	501	
Expected effect of treatment on employment prospects			
improve strongly	0.27	0.54	0.00
improve somewhat	0.40	0.37	0.44
unchanged	0.31	0.08	0.00
worsen somewhat	0.02	0.01	0.19
worsen strongly	0.01	0.00	0.15
B. Short-term training			
No. of observations	486	971	
improve strongly	0.25	0.34	0.00
improve somewhat	0.41	0.45	0.11
unchanged	0.30	0.18	0.00
worsen somewhat	0.02	0.02	0.78
worsen strongly	0.02	0.01	0.11
C. Non-participants			
No. of observations	2,163	2,138	
improve strongly	0.15	0.38	0.00
improve somewhat	0.43	0.46	0.02
unchanged	0.39	0.14	0.00
worsen somewhat	0.01	0.01	0.10
worsen strongly	0.02	0.01	0.00

Note: Percentage share unless indicated otherwise. $P-$values measured based on two-tailed t-tests on equal means.

Table B.2: Marginal Effects of Probit Estimation: Selection into Realized Treatment (Full Estimates)

	Long-term training v. non-participation		Short-term training v. non-participation		Long-term training v. short-term training	
	Coef.	SE	Coef.	SE	Coef.	SE
Expected participation probability: high ($\pi \geq 5$)	0.0826***	(0.0125)	0.0865***	(0.0129)	0.0257	(0.0216)
Expected effect of treatment on employment prospects						
Improve strongly	0.1123***	(0.0186)	0.0221	(0.0161)	0.1820***	(0.0302)
Improve somewhat	0.0297**	(0.0141)	-0.0027	(0.0141)	0.0607**	(0.0275)
Female	0.0250**	(0.0120)	0.0101	(0.0135)	0.0335	(0.0234)
Age (ref.: 16-24 years)						
25-34 years	0.0445**	(0.0178)	-0.0322	(0.0203)	0.1080***	(0.0364)
35-44 years	0.0682***	(0.0231)	-0.0416*	(0.0238)	0.1540***	(0.0447)
45-5 years	0.0741***	(0.0249)	-0.0586**	(0.0238)	0.1651***	(0.0470)
School leaving degree (Ref.: None)						
Lower sec. degree	0.0197	(0.0338)	-0.0370	(0.0361)	0.0918	(0.0644)
Middle sec. degree	0.0364	(0.0368)	-0.0354	(0.0365)	0.1262*	(0.0678)
Upper sec. degree	0.0520	(0.0406)	-0.0857**	(0.0343)	0.2031***	(0.0766)
Higher education (Ref.: None)						
Internal/external prof. training	-0.0029	(0.0180)	0.0438**	(0.0209)	-0.0606*	(0.0357)
University degree	-0.0081	(0.0213)	0.0058	(0.0247)	-0.0246	(0.0466)
German citizenship	0.0216	(0.0267)	-0.0338	(0.0309)	0.0761	(0.0536)
Migration background	0.0064	(0.0171)	-0.0259	(0.0196)	0.0291	(0.0353)
Married or cohabiting	-0.0130	(0.0117)	0.0144	(0.0150)	-0.0354	(0.0235)
Problems with childcare	-0.0069	(0.0186)	0.0130	(0.0239)	-0.0167	(0.0375)
Partner is full-time employed	-0.0295	(0.0195)	-0.0318	(0.0245)	-0.0199	(0.0491)
Children in household						
Age 0-3 years	0.0232	(0.0195)	0.0061	(0.0221)	0.0321	(0.0370)
Age 4-6 years	0.0261	(0.0206)	-0.0106	(0.0238)	0.0571	(0.0422)
Age 7-15 years	0.0146	(0.0147)	0.0071	(0.0173)	0.0190	(0.0286)
Age 16-18 years	0.0271	(0.0193)	0.0192	(0.0225)	0.0229	(0.0357)
Searching for full-time employment	-0.0300**	(0.0147)	-0.0299	(0.0190)	-0.0405	(0.0328)
Region (ref.: West-Germany & UE rate 0-6%)						
West-Germany & UE rate 6%+	-0.0039	(0.0116)	-0.0315**	(0.0136)	0.0223	(0.0235)
East-Germany & UE rate 9-14%	-0.0184	(0.0147)	-0.0190	(0.0182)	-0.0035	(0.0307)
East-Germany & UE rate 14%+	0.0069	(0.0167)	-0.0533***	(0.0180)	0.0763**	(0.0349)
Entry into unemployment (ref. 2nd quarter 2007)						
3rd quarter 2007	-0.0242	(0.0190)	-0.0315	(0.0240)	-0.0183	(0.0415)
4th quarter 2007	-0.0013	(0.0204)	-0.0055	(0.0247)	0.0016	(0.0407)
1st quarter 2008	-0.0116	(0.0218)	0.0011	(0.0277)	-0.0296	(0.0428)
2nd quarter 2008	0.0146	(0.0234)	-0.0375	(0.0253)	0.0644	(0.0468)
Time to interview (ref.: 7 weeks)						
8 weeks	0.0155	(0.0402)	-0.0063	(0.0426)	0.0445	(0.0837)
9 weeks	0.0220	(0.0423)	0.0313	(0.0471)	0.0185	(0.0822)
10 weeks	0.0512	(0.0488)	0.0195	(0.0472)	0.0714	(0.0892)
11 weeks	0.0320	(0.0458)	0.0189	(0.0481)	0.0266	(0.0862)
12 weeks	0.0253	(0.0462)	0.0134	(0.0496)	0.0333	(0.0908)
13 weeks	0.0153	(0.0477)	-0.0133	(0.0516)	0.0693	(0.104)
14 weeks	0.0168	(0.0459)	0.0227	(0.0523)	-0.0002	(0.0890)
Unemployment benefit recipient	0.0106	(0.0134)	0.0503***	(0.0162)	-0.0334	(0.0271)
Last daily income in €	0.0002	(0.0002)	-0.0005*	(0.0002)	0.0011***	(0.0004)
Employment status before unemployment (Ref.: Other)						
Regular employment	0.0081	(0.0191)	0.0300	(0.0230)	-0.0203	(0.0388)
Subsidized employment	0.0112	(0.0223)	0.0122	(0.0254)	-0.0010	(0.0458)
Last job was full-time employment	-0.0104	(0.0122)	-0.0240*	(0.0143)	0.0066	(0.0257)
Months in employment						
in last year	-0.0012	(0.0015)	-0.0013	(0.0018)	-0.0028	(0.0031)
in last 5 years	0.0004	(0.0006)	0.0009	(0.0007)	-0.0001	(0.0013)
in last 10 years	0.0001	(0.0003)	0.0003	(0.0004)	-0.0002	(0.0007)
Months in unemployment						
in last year	-0.0026	(0.0033)	-0.0089**	(0.0040)	0.0046	(0.0067)
in last 5 years	0.0017	(0.0012)	-0.0035**	(0.0014)	0.0062***	(0.0024)
in last 10 years	-0.0007	(0.0006)	0.0015**	(0.0007)	-0.0031***	(0.0012)
Openness	-0.0072	(0.0053)	-0.0037	(0.0062)	-0.0048	(0.0105)
Conscientiousness	0.0097*	(0.0056)	0.0008	(0.0063)	0.0157	(0.0114)
Extraversion	-0.0207***	(0.0056)	-0.0058	(0.0065)	-0.0282**	(0.0110)
Neuroticism	-0.0045	(0.0053)	0.0030	(0.0063)	-0.0135	(0.0109)
Locus of control	0.0006	(0.0053)	-0.0180***	(0.0064)	0.0165	(0.0111)
No. of observations	5,008		5,758		2,164	
log-Likelihood	-1886.6		-3137.4		-1266.4	
$P-$value (LR-test)	0.0000		0.0000		0.0000	
Pseudo-R^2	0.0746		0.0367		0.0738	

Note: Depicted are average marginal effects based on probit models estimating the realized participation probability. Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level.

Table B.3: Sensitivity Analysis: Difference-in-Difference Estimates for Alternative Expectation Measure

	Effect of expected participation probability π-high v. π-low		
	Long-term training	Short-term training	Difference-in-Difference[(a)]
	IPW	IPW	DID
	(1)	(2)	(3)
Regular employed in month			
$t+12$	0.1109***	-0.0978***	0.2087***
	(0.0404)	(0.0293)	(0.0499)
$t+30$	0.0953**	-0.0065	0.1019*
	(0.0447)	(0.0290)	(0.0533)
Average monthly earnings in €	-67.69	-82.03	14.33
	(170.74)	(66.81)	(183.27)
No. of observations	707	1,457	2,164
Control variables			
Socio-demographic characteristics	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes
Labor market histories	Yes	Yes	Yes
Regional and seasonal information	Yes	Yes	Yes
Personality traits	Yes	Yes	Yes

Note: Depicted are average treatment effects of the expected participation probability π-high (relative to π-low) based on inverse probability weighting (IPW) using an alternative expectation measure that refers to the expected probability to participate in any labor market program. Standard errors are shown in parenthesis. */**/*** indicates statistical significance at the 10%/5%/1%-level. [(a)]The difference-in-difference model (column 3) shows the impact of π-high on participants in long-term training relative to the impact of π-high on participants in short-term training.

Figure B.1: Empirical Setting and Economic Framework

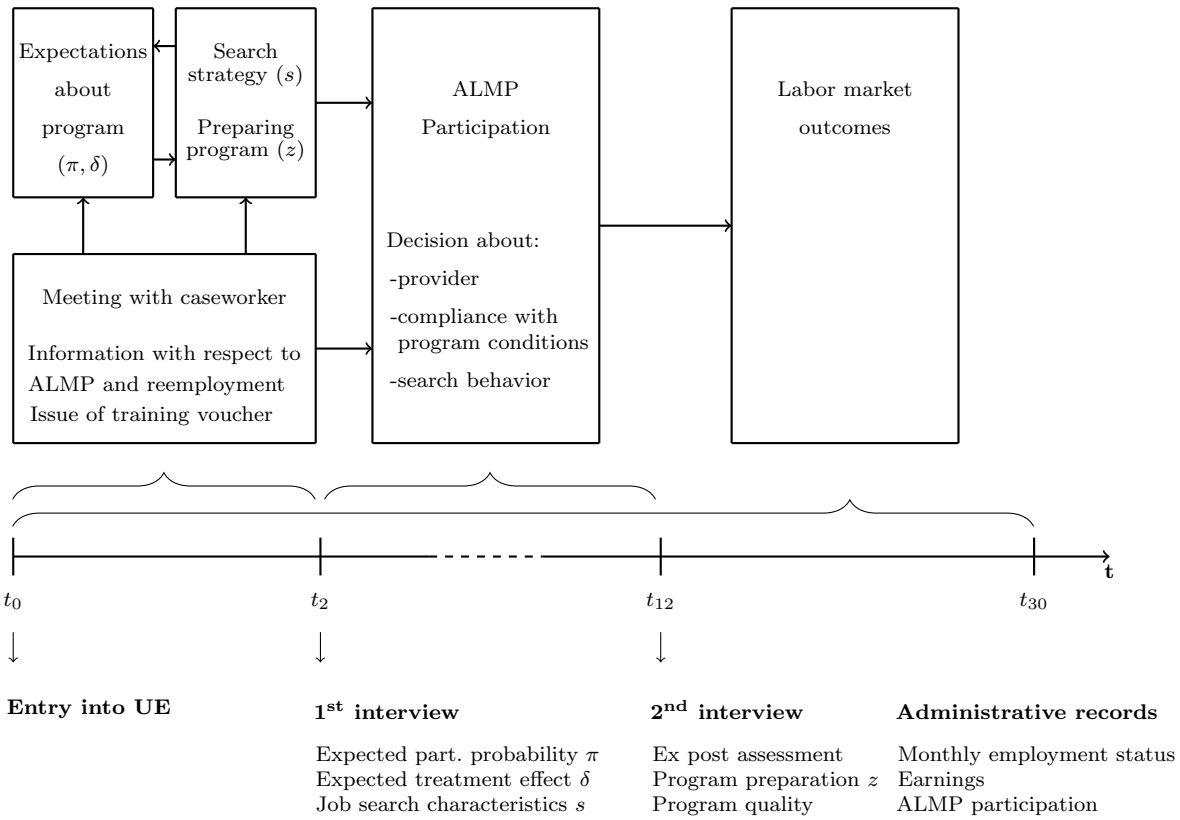
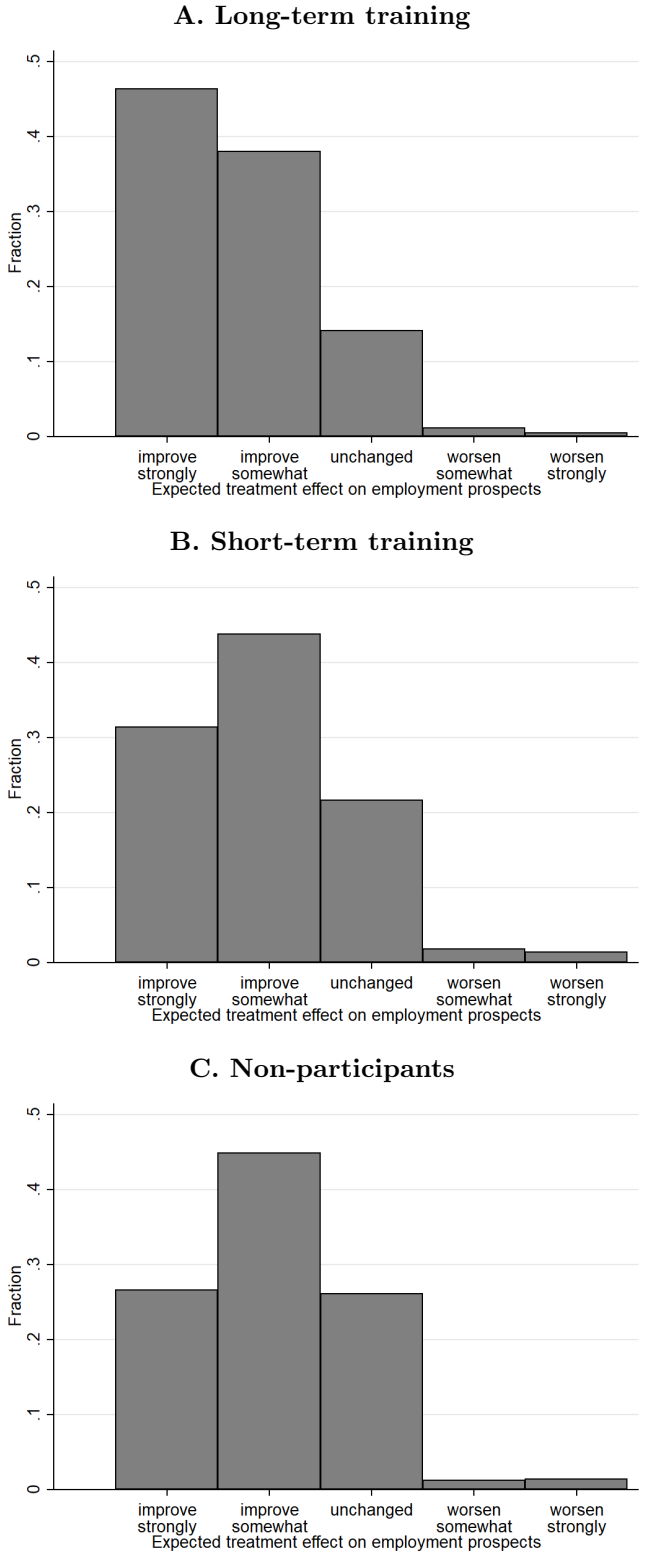
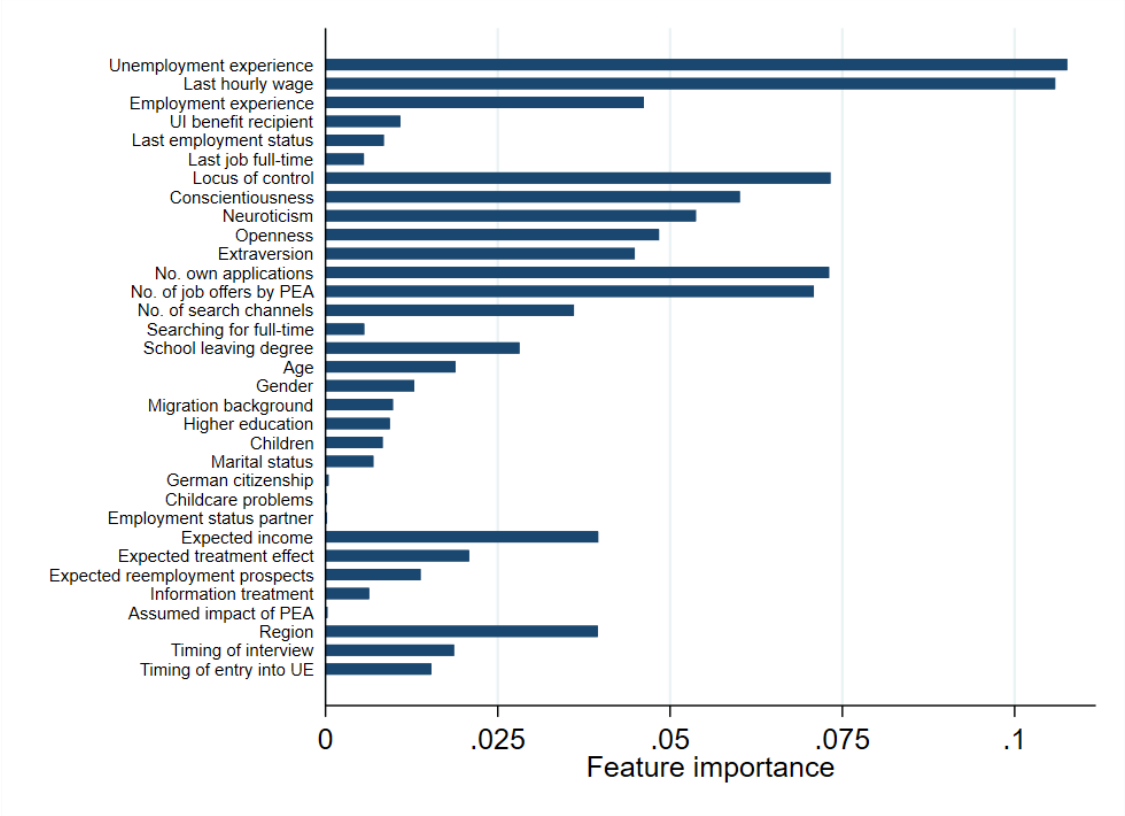


Figure B.2: Distribution of the Expected Treatment Effect



Depicted are answers to the question: *"In your opinion, to what extent would your chances of finding new employment be changed by participation in a training scheme?"* 1 = improve strongly, 2 = improve somewhat, 3 = remain unchanged, 4 = worsen somewhat, 5 = worsen strongly.

Figure B.3: Feature Importance for Baseline Specification



Note: Feature importance is defined as the normalized average reduction in the expected mean squared error after including a specific variable. Variables are listed in descending order of feature importance. The baseline specification includes the full set of variables. The most important features within each category enter the final specification (see Figure 3).

Figure B.4: Distribution of Conditional Average Treatment Effects