

Brändle, Tobias; Fervers, Lukas

Article — Published Version

Give it Another Try: What are the Effects of a Job Creation Scheme Especially Designed for Hard-to-Place Workers?

Journal of Labor Research

Provided in Cooperation with:

Springer Nature

Suggested Citation: Brändle, Tobias; Fervers, Lukas (2021) : Give it Another Try: What are the Effects of a Job Creation Scheme Especially Designed for Hard-to-Place Workers?, Journal of Labor Research, ISSN 1936-4768, Springer US, New York, NY, Vol. 42, Iss. 3-4, pp. 382-417, <https://doi.org/10.1007/s12122-021-09322-x>

This Version is available at:

<https://hdl.handle.net/10419/287139>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Give it Another Try: What are the Effects of a Job Creation Scheme Especially Designed for Hard-to-Place Workers?

Tobias Brändle¹ · Lukas Fervers²

Accepted: 15 September 2021 / Published online: 2 October 2021
© The Author(s) 2021

Abstract

Previous evaluations of job creation schemes (JCS) reveal mostly negative employment effects, mainly due to inherent lock-in effects. In this paper, we assess the impact of an innovative JCS that employs a pre-selection mechanism to target programme participation on unemployed job seekers with very low integration chances, hereby reducing possible lock-in effects. Relying on high-quality administrative as well as survey data, we conduct regression-adjusted matching analyses to estimate the programme effect on integration into regular employment. Our results show that the programme did not succeed to foster labour market integration, but still entails remarkably negative employment effects in the first years after participation. We argue that this results from a principal-agent problem at the last step of the selection mechanism that may have led to cream-skimming rather than targeting on very hard to place workers. However, supplementary analyses reveal that negative effects can be avoided for subgroups with very poor employment chances in case of non-participation. These results are robust to the use of different matching estimators and definitions of non-participation. The inclusion of usually unobservable survey variables as well as placebo tests based on past employment outcomes refute concerns about endogenous selection. From a policy-perspective, these findings imply that targeting JCS on workers with very low integration chances is a key factor to avoid negative employment effects found in previous evaluations. At the methodological level, our analyses add to recent literature that assesses the credibility of non-experimental evaluations based on high-quality administrative data.

✉ Lukas Fervers
fervers@wiso.uni-koeln.de
Tobias Brändle
Tobias.Braendle@iaw.edu

¹ Institute for Applied Economic Research (IAW), Schaffhausenstraße 73, 72072 Tübingen, Germany

² University of Cologne, German Institute for Adult Education (DIE) and Institute for Applied Economic Research (IAW), Albertus-Magnus Platz, 50923 Cologne, Germany

Introduction

Since the mid-2000s, Germany has experienced a remarkable improvement in labour market performance. After having reached its peak of 11.2% in 2005, the unemployment rate has fallen to 4.6% within the following ten years, even initiating a discussion on a new German miracle (Rinne and Zimmermann 2012; Caliendo and Hogenacker 2012; Möller 2015). At the same time, despite more than half a million of registered job vacancies (in 2015; OECD 2020), the number of (very-) long-term unemployed workers remains rather high (Spermann 2015). Apparently, a certain share of the long-term unemployed faces persistent difficulties in meeting the demands of employers, even under favourable market conditions and many open positions.¹

Among others, this mismatch may result from a depreciation of skills during long-term unemployment, increasing behavioural problems, or discouragement after long periods of unsuccessful job search (DeLoach and Kurt 2013). In such dire circumstances, active labour market policy seems to be an appropriate tool to bringing these individuals back to the labour market. Consequently, public job creation schemes (JCS) have been heavily used in the past to restore basic cognitive and behavioural skills of long-term unemployed workers. Providing basic skills required in the labour market, they were supposed to function as a stepping-stone into regular employment. However, previous research shows that JCS mostly fail to reach their goal and even tend to worsen the chances to find regular employment (Thomsen and Walter 2010; Hujer and Thomsen 2010; Wolff and Stephan 2013, for reviews see Card et al. 2010 and Kluge 2010). While directly assessing the mechanisms behind these negative effects is challenging, two frequently suggested explanations refer to the low skill intensity of the JCS jobs as well as possible lock-in effects. As participants often stop searching for regular employment, integration rates into regular employment are lower compared to the hypothetical scenario of non-participation.

To address this inherent shortcoming of former JCS, Germany has initiated an innovative programme aimed at increasing the acquisition of skills during the JCS and preventing lock-in effects. The key idea is the use of a special selection mechanism when choosing participants. Before becoming eligible for the JCS jobs, programme participants undergo a period of intensified counselling and monitoring that lasts for at least six months. Only those who cannot find a job within this period are allowed to apply for the JCS jobs. Moreover, participation has been restricted to unemployed workers who have received social assistance for at least one year before programme start. In most cases, this implies a period of two years of unemployment. This approach follows a rather simple yet (seemingly) convincing idea: By targeting the programme on very long-term unemployed workers and filtering those who

¹ The comparatively high share of long-term unemployed workers has been a reoccurring pattern in the German labour market. It has been attenuated only temporarily in the second half of the last decade (2015–2020). During the Corona crisis, the number of long-term unemployed job seekers has jumped back quickly to over one million. The share of unemployed job seekers amounts to 35.1% in 2021 (Statista 2021).

can find a job with more intense support, those who enter the JCS jobs would have very low integration rates in case of non-participation anyway. Consequently, lock-in effects lose their relevance and the programme is likely to have more favourable effects. In other words: If counterfactual integration rates in case of non-participation are very low, the programme effect cannot be negative anymore.

In this paper, we wish to assess whether this innovative approach succeeds to exert a positive influence on labour market reintegration. To this end, we follow previous evaluation research and employ a control group design to estimate the impact of the programme on labour market integration. The control group consists of persons who have undergone the activation period but did not participate in the JCS afterwards. From a methodological point of view, this creates rather favourable conditions for non-experimental evaluation because treatment and control group have undergone the same pre-selection mechanism. Therefore, they can be assumed to be rather similar to each other. Relying on high-quality register data, we perform regression-adjusted matching/weighting analyses to estimate the average treatment effect on the treated. To substantiate the validity and robustness of our results, we rerun the analysis with additional (usually unobservable) variables on behavioural and psycho-social characteristics. Moreover, we employ placebo tests based on past employment outcomes (Heckman and Hotz 1989; Imbens and Rubin 2015).

Despite the innovative institutional setting, our results show that the programme did again not succeed in improving employment chances. In contrast, programme effects remain remarkably negative and reduce the probability to find regular employment by up to 50% in the first years after programme start. These results are robust to different methodological approaches (e.g. different matching/weighting algorithms, different definitions of non-participation) and rarely change when (usually unobservable) survey variables are included in the analysis. We argue that this negative effect results from a principal agent problem at the last stage of the selection mechanism. As JCS employers could select between different candidates, the last step entails cream-skimming rather than effective targeting on unemployed workers with little employment chances. At the same time, supplementary analyses show that the effect is at least not negative for those participants who are in a particularly disadvantaged labour market position. In sum, the idea of effective targeting remains essential for more favourable programme effects, but the selection mechanism employed here has not been rigorous enough. These results add to the growing literature on policy measures for long-term unemployed workers in general (Card et al. 2010; Kluve 2010; Arni et al. 2013; Arendt and Kołodziejczyk 2019) and the impact of JCS in particular (Caliendo et al. 2004, 2005, 2008; Lechner and Wunsch 2009; Hujer and Thomsen 2010). At the methodological level, our results complement previous work that assesses the validity of non-experimental evaluations based on high quality register data (Lechner and Wunsch 2013; Caliendo et al. 2017).

The remainder of the paper is organized as follows. We start by briefly reviewing previous work on ALMP programmes, especially JCS. Subsequently, we explain the institutional setting in more detail before we outline our empirical analyses. The final section concludes with a short summary of the results and discussion on the implications for policy-making and future research.

Literature Review

There has been an increasing interest in the impact of JCSs since the end of the 1990s/ beginning of 2000s. Most of the early work focusses on East Germany where JCS have been heavily employed as secondary labour markets after German reunification. Overall, the results are rather negative, programme participation even reduces the chances of finding regular employment. Based on an administrative dataset, Lechner and Wunsch (2009) and Kraus et al. (2004) analyse different JCSs from East Germany and find negative effects for all programmes. Further research on East and West Germany conducted by Caliendo et al. (2004, 2005, and 2008), Hujer and Thomsen (2010) and Heyer et al. (2012) arrives at the same conclusion. These microeconomic studies are complemented by a macroeconomic evaluation conducted by Hujer and Zeiss (2005). The estimated augmented matching function confirms a negative impact of an increasing inflow into JCSs on inflow into regular employment at the aggregate level. Outside the German context, JCSs have gained increasing importance in Switzerland towards the end of the 1990s due to an exceptionally long period of economic stagnation and rising unemployment. Gerfin and Lechner (2002) rely on administrative data and provide empirical evidence for a wide range of different ALMPs. While the overall results are quite mixed, JCSs are shown to insert a consistently negative effect on employment probability within the first year after programme start. In Sweden, ALMPs have always played an important role, regardless of the current economic situation (for a comparative evaluation of different programmes see Frölich et al. 2004 or Carling and Richardson 2004). Even though the economic and institutional context strongly differs from the one in the aforementioned studies, the results (e.g. Sianesi 2008) again point to negative employment effects. Apparently, the negative effect of JCS is a consistent finding regardless of the economic circumstances. This general notion is confirmed by meta-analyses conducted by Card et al. (2010), Kluge (2010) as well as Card et al. (2018).

These negative average effects raise the question of whether JCSs may at least be beneficial for certain groups. Caliendo et al. (2004, 2005, 2008) have made a start and distinguished the effects with respect to the usual suspects for effect heterogeneity, namely region (East vs. West Germany), gender, and sector of employment. While they do find significant differences for some sub-groups, there is no clear-cut pattern across all analyses. A less ambiguous picture is revealed by Caliendo et al. (2008) as well as Hujer and Thomsen (2010) with respect to foregoing unemployment duration. Hujer and Thomsen (2010) stratify their sample according to the number of quarters of unemployment and show that the effect is less negative for persons with longer previous unemployment duration. This supports the argument that programmes are at least less detrimental to persons with lower labour market attachment, as lock-in effects are less relevant for this group.

Finally, it is worth looking at the magnitude and the temporal pattern of the effects. The latter are typically negative right after programme start, and accelerate up to a certain point in time, but get smaller (or sometimes even insignificant)

towards the end or after the programme. The importance of lock-in effects becomes apparent when comparing JCSs to activation programmes that have a similar content but are less encompassing. Evaluations of the so-called One-Euro jobs (a German workfare programme which offers temporary employment for a short period of time (3–12 months) in exchange for low monetary compensation) point to positive or at least mixed effects on re-employment chances (Huber et al. 2011; Hohmeyer 2012; Hohmeyer and Wolff 2012; Dengler 2015). At the same time, it should be noted that negative effects of JCS often persist after the end of the programme. Hujer and Thomsen (2010) analyse JCSs which typically last for 12 months. They report a reduced employment probability of 9% points even two years after programme start for short-term unemployed males. Similarly, Sianesi (2008, p. 386) reports negative effects even five years after programme start. Card et al. (2018) suggest that the slow recovery from the lock-in period may come from the lack of appreciation of private employers to the experience or skills gained in JCSs. This argument is supported by the more positive findings concerning supplementary jobs in the regular labour market, which entail more intense human capital accumulation as well as more positive signalling effects (Mosthaf et al. 2021). The detected impacts of JCS are often large in magnitude. Hujer and Thomsen (2010, p. 45) report negative employment effects of 20.8% points six months after programme start (women: 28.8% points). Even for male workers who enter the JCS during their fifth quarter of unemployment, the effects amount to 15.8% points (women: 22.1% points). Results from other countries fall in a similar range. Sianesi (2008, p. 386) reports negative short-term average effects of more than 20% points, Lechner and Wunsch (2009, p. 685) estimate a negative effect of about 25% points six months after programme start.

In sum, three conclusions can be drawn. First, JCSs tend to have large adverse effects on employment outcomes for participants. Second, the effects are very strong immediately after programme start but get weaker towards the end of the observation period. Finally, the effects differ between groups, especially with respect to foregoing unemployment duration. However, it is important to note that the short-term effects are still quite large, even for long-term unemployed workers. This implies that the counterfactual employment probabilities in case of non-participation would not have been that low. JCSs might therefore be more effective when targeted on individuals with very low employment chances. Yet, previous evidence suggests that targeting simply based on foregoing unemployment duration as the main or only criterion of selection might not be sufficient to identify workers with low employment chances. This raises the question whether there could be alternative selection mechanisms that can identify more suitable target groups.

Structure of the Programme

Following past experiences, Germany has initiated a new JCS (called “Bürgerarbeit”) with some distinct features aimed at overcoming the shortcomings of previous approaches. The programme under discussion has been run in Germany between 2011 and 2014 (the vast majority of entries into the JCS happened between January

2011 and mid of 2012) with participating regions scattered throughout the country. Similar to previous JCS, its basic idea is that unemployed workers get a publicly financed job that has to be of public utility. These jobs are mostly located in the public sector or at charity organizations and must not substitute regular employment. The activities carried out range from social services (e.g. transport services for charity organizations) to manual occupations or administrative tasks (for a detailed description of the tasks see IAW and ISG 2015). The programme has offered 33,955 JCS jobs to unemployed job seekers who have received social assistance for at least one year. This restriction implies that potential programme participants have mostly been unemployed for at least two years. Consequently, potential programme participants consist of long-term unemployed job seekers with particularly unfavourable employment histories. The JCS constitutes a regular employment contract subject to social security contributions with a gross monthly wage of at least 900 Euros (30 h/week, 600 Euros in case of 20 working-hours/week). The income can be combined with social assistance if total household income remains below the social assistance threshold. In case of a 30-h contract, total household income usually strongly exceeds the one based exclusively on welfare benefit receipt. The duration of the JCS ranges between one and three years.

Apart from these basic characteristics, there are two special features that make the programme particularly interesting. First, the programme offers part-time rather than full-time jobs. This guarantees that participants have sufficient time for job search in the regular labour market. In addition, participants receive job-related, individualized coaching. The coaching may cover personal problems as well as further training for skills related to the labour market. Second and more importantly, selection into the programme follows a special mechanism. Before participants can apply for the JCS, they have to undergo a period of intensified counselling and monitoring lasting at least for six months (*activation period*), in which they have to search for a job in the regular labour market. During the activation period, they receive special support by the local employment agencies that goes beyond basic services. Only if they cannot find a job within the activation period, they can apply for the JCS. Those who cannot find a JCS job continue to look for a job in the regular labour market without special support. The idea of this pre-selection mechanism can easily be explained in the spirit of the potential outcome framework (for a formal outline see subsequent section): The causal effect of the programme is equal to the difference between the outcome in case of participation and the hypothetical outcome in case of non-participation. As previous research has shown, the (estimated) hypothetical outcome in case of non-participation was not that low, leading to negative treatment effects due to lock-in effects. Consequently, the programme might be more effective if it is targeted on unemployed job seekers with very low or even no employment chances in case of non-participation. As outlined in the literature review, previous research has established that targeting based on observable characteristics such as region of residence or employment history alone is insufficient to avoid negative treatment effects. The outlined pre-selection via the activation period can therefore be regarded as a new attempt to target the JCS on workers with very low employment chances. On the one hand, this avoids lock-in effects. On the other hand, this group is likely to profit more from the basic skills offered by the programme. Previous research on

the activation period has confirmed that it indeed fosters exit into employment, which means that it successfully filters out individuals who could find a job in a relatively short period of time using increased activation and monitoring measures only. In total, about 10% of participants at the activation period have found employment before the end of the activation period, which is 2.5% points higher than the matched control group (IAW and ISG 2015; Fervers 2019).

While the institutional setting of the programme seems to be a promising approach to tackle the shortcomings identified for JCS in the past, it should be noted that its idea is partly undermined by the final step in the selection process of the programme. In most cases, more than one (on average three) job seekers apply and compete for one JCS job. The final hiring decision is made by the JCS-employer. We argue that this mechanism essentially creates a principal-agent problem. Even though JCS-employers may differ in their utility function from private employers, with e.g. productivity of the employees being less important, they are still likely to pick individuals with favourable personal characteristics and character traits. In contrast to the idea of the programme, this suggests that JCS-employers might not pick applicants with particularly severe placement obstacles. In short, this mechanism entails cream-skimming rather than targeting on hard-to-place workers.

At the methodological level, this creates additional challenges because it gives rise to systematic selection. Given that employers select between different candidates based on job interviews, they are likely to observe and base their decision on characteristics such as motivation or communication skills. As these characteristics are unobserved in administrative data, this might lead to endogenous selection. At the same time, it should be considered that the institutional setting is rather favourable for non-experimental evaluation, as the group that is split into treatment and control group consists of rather homogenous individuals. Selection into the JCS after the activation period only occurs in the advertisement for the JCS jobs. Whether the unemployed job seekers choose to apply for a JCS job is likely to depend on several exogenous factors that are unrelated to his or her (possibly unobserved) characteristics. These include the overall availability or number of JCS-jobs in the particular region or the fit between the available jobs and his or her job preferences. Nevertheless, we will carefully assess the CIA by both placebo tests as well as the inclusion of (usually unobservable) control variables that capture unobserved character traits (see robustness section).

Empirical Analysis

Research Design, Data and Estimation

Research Design

Following previous econometric evaluation research, we rely on a control group design to estimate the treatment effect of the programme. In the terminology of the potential outcome framework (Rubin 1974; Rubin and Imbens, 2015), the treatment effect on the treated τ_{ATT} can be defined as the difference between the

outcome that has been realized in case of participation (Y^1) and the hypothetical outcome in case of non-participation (Y^0). In expectations, this can be expressed as $\tau_{ATT} = E[Y_i^1 - Y_i^0 | D = 1]$. As Y_i^0 cannot be observed, a control group is used to estimate the treatment effect by comparing observed outcomes of the treatment and control group. Estimating the treatment effect by comparing observed outcomes hinges on the assumption that potential outcomes of treatment and control group would be the same under a certain treatment condition, i.e. that treatment status is independent from potential outcomes ($D_i \perp Y_i^0, Y^1$). As we rely on a rich set of covariates, this assumption can be relaxed to the conditional independence assumption (CIA; $D_i \perp Y_i^0, Y^1 | X_i$). Applying this framework to our research question, this implies that we compare the employment outcomes of the treatment and a control group after treatment start. Doing so, we assume that any selection into the programme is due to observable variables. Considering the institutional setting, this assumption is far from trivial. Given that JCS employers pick between different candidates, it is rather likely that there will be selection on observables, but possibly also on unobservables. We are therefore well advised to assess the credibility of the CIA by all possible means (see section on estimation).

Data and Variables

The most important data sources are the Integrated Employment Biographies (IEB) that consist of the social security records of participants. The data contain detailed information on past employment biographies. This includes all spells of employment, unemployment and programme participation. The IEB are therefore commonly used in labour market research in the German context (for a detailed description see Dorner et al. 2010 and Biewen et al. 2014). In total, we have access to a random sample of 63,743 individuals who participated in the activation period but were still unemployed at its end. 12,207 of these individuals became employed in the JCS (treatment group), the remaining 51,536 individuals kept looking for a job without special support (control group). From the social security records, we code dummy outcome variables that indicate whether someone has been employed at a certain point in time after the (non-) treatment start (coded in 30 days intervals). Employment is defined as having an unsubsidized job in the regular labour market, i.e. a job that entails social security contributions but is not part of another ALMP scheme. In addition, the social security records contain a rich set of information that are used as covariates. The first group of covariates includes basic socio-demographic information such as education, age, family status, foreign nationality, number of persons in the household, having children in a certain age and having health problems. Second and most importantly, we count the number of months in the respective employment state in the last, second to fourth, and fifth to seventh year before programme start plus a dummy variable indicating whether someone has ever been employed during the last seven years. Finally, we include information on certain aspects of previous employment and programme history such as complexity of last job, blue vs. white collar worker, sector of the industry, drop out from/successful completion

of an ALMP programme and the subjective assessment of the counsellor in the employment agency on future employment prospects. In addition to the individual-level variables, we rely on official statistics on the local unemployment rate, employment rate, GDP and population density on the county level. All in all, this data reflects the state-of-the-art in non-experimental evaluation studies (Lechner and Wunsch 2013).

Due to the institutional setting, one might still be concerned that register data alone may be insufficient. Given the assignment mechanism, Caliendo et al. (2017) have shown selection into ALMPs might also be due to variables such as personal reliability, communication skills, motivation or related aspects that are observable in a job interview but not included in the register data. Therefore, we have conducted a short survey for potential participants in the JCS, i.e. individuals at the end of the activation period. Survey variables include self-assessed personal skills, particular social-psychological problems, application behaviour and willingness to make concessions to find a job (e.g. acceptance of commuting time or shift work) and support by institutions or the peer group (family and friends). These variables serve as measurement for individual motivation, search intensity, existence of further placement obstacles, and personality. To limit the number of variables included in the propensity score estimation, we have used principal component analyses to reduce the item batteries used in the survey into one or two indicator variables each. The number of resulting indicators is determined by the value of the eigenvalues.

Finally, we have coded treatment status in two different ways, the dynamic matching (see Biewen et al. 2014) and the hypothetical start date approach (Lechner 1999). The definition of non-treatment is a non-trivial issue in two regards. First, if treatment assignment is dynamic rather than static, it is controversial how units of observation that are not treated at the very beginning shall be dealt. Second, if one limits the control group to persons who never participate, the start of non-treatment has to be defined. In the econometric literature, two approaches have been developed to address these problems. The hypothetical start date approach restricts the control group to persons who never participate. While this avoids the problem of treated persons becoming part of the control group, control observations still have to be assigned a start date of non-treatment. To keep the temporal pattern between treatment and control group symmetrical, we have generated a random variable that mirrors the transition pattern of the treatment group from the activation period to the start of the JCS job. We hereby make sure that—at the start of treatment/non-treatment—treatment and control observations have spent the same amount of time being at risk before treatment start. The distribution of the random variable is given in Fig. 6. While this approach has the advantage of comparing treated units to units who do not receive the same treatment at any point in time, it has been criticized for conditioning on future outcomes. In our setting, this might entail a downward bias of the treatment effect estimates as the chance of never being treated in the future increases if someone finds a job in the regular labour market. Therefore, the dynamic matching approach proceeds in a different way and compares all persons who start the treatment in a certain period of time to all persons who did not, regardless of their treatment status in the future (the not-yet treated, see Dauth and Toomet, 2016). We therefore stratify our sample in monthly intervals. Those who start the

treatment in a certain month are defined as treatment group, all eligible unemployed job seekers who did not start the treatment in the respective month are defined as control group. In the estimation, all monthly samples are pooled together.

A complete list of all variables including summary statistics is outlined in Tables 1 (register variables) and 2 (survey variables). Overall, the descriptives reinforce that the group of unemployed job seekers are characterized by rather low labour market attachment. This results from the pre-selection due to the activation period as well as the selection into the activation period, as individuals with rather unfavourable characteristics have participated here (see Fervers 2019).

Estimation

We apply regression-adjusted matching estimations. For sake of robustness, we employ two different matching estimators that have performed well in previous validation studies, namely radius matching as suggested by Huber et al. (2013) and entropy balancing developed by Hainmueller (2012). For both algorithms, we check the matching quality by assessing standardized bias for all covariates before and after matching. As the results from the propensity score estimations (see Table 3) suggest, there is some but no strong covariate imbalance before matching. Participants tend to be older, have a higher level of education, are less often foreign nationals and come more frequently from the medium categories of the subjective placement assessment by the caseworkers. However, covariate imbalance is limited (propensity score estimation pseudo $R^2=0.16$), which (together with the high number of available control units) refutes concerns about thin common support. Correspondingly, both matching algorithms succeed to strongly minimize covariate imbalance after matching to a sufficient degree. For radius matching, mean bias is reduced from 10.8 to 3.9 in the baseline specification (median bias from 7.8 to 3.0; for a summary of standardized bias before and after matching see Fig. 7). Given the high number of observations and weak to medium selectivity of the treatment, it comes as no surprise that differences in the results between both estimators are limited in all specifications.

While checking matching quality is straightforward, assessing the credibility of the underlying assumptions is more difficult. Since the CIA is fundamentally untestable, we proceed in three ways to test its credibility in an indirect way. First, we follow the approach suggested by Caliendo et al. (2017) and conduct the analysis twice, with and without survey variables on motivation, search intensity and personality. As the survey variables are only available for a subsample, we check whether possible differences could be due to sample selection by repeating the analysis a third time without survey variables but only with observations for whom survey variables are available. The idea behind this approach is to check the sensitivity of the results for usually unobservable variables after conditioning on information from the register data. If results rarely change, it seems more reasonable to argue that the results would not change if even more variables are conditioned on, either. This implies that a hidden selection bias is less likely. Second, we follow the approach suggested by Heckman and Hotz (1989) and reinforced by Rubin and Imbens (2015) and perform a series of placebo tests on past employment outcomes. To this end, we define the number of months in employment in the time span ten to nine, nine to

eight and eight to seven years before programme start as placebo outcome variable, respectively (these variables are not used as covariates in the matching analysis). We then repeat our analyses with the placebo outcomes as dependent variables. Significant effects would point to persistent unobserved differences between groups, as the placebo effects cannot be explained by a treatment that has taken place several years later. We conduct the placebo tests for the estimations with and without the survey variables with both outlined matching algorithms. Third, we conduct a series of estimations that relax the CIA when external instruments are not available. These estimation techniques have been developed and discussed in Chen and Wang (2020), Millimet and Tchernis (2013) as well as McCarthy et al. (2014).

Finally, one might worry about interference between units. The stable unit value assumption (SUTVA) might be violated in case of substitution effects when treated workers replace non-treated workers. However, the SUTVA has already been assessed in previous evaluations of the activation period by comparing employment outcomes from unemployed non-participants in participating and non-participating regions in a difference-in-differences framework (Fervers 2019). The estimated substitution effect is almost zero and insignificant for all points in time (including time spans in which the JCS has already started) after programme implementation. It should be noted that this test has a very high statistical power as the number of observations exceeds 200,000. We therefore feel safe in assuming that substitution effects play a minor role in this programme.

Results

Figure 1 summarizes the results of the treatment effect estimations based on radius matching (left panel) and entropy balancing (right panel). The lower panel shows the results for dynamic matching. The treatment effect gets negative right after the start of the programme and accelerates during the first 18 months. In absolute terms, the treatment effect estimated by the hypothetical treatment start date approach reaches up to -9.2% points. Given that absolute integration rates of the matched control group remain below 20% throughout the observation period, this translates into a remarkably strong relative effect. Considering that programme duration ranges between one and three years, the observation period can at least be regarded as medium-run analysis. It should be noted that not all persons can be observed until the end of the observation period. Since some observations are censored after more than 18 months, the long-run effects have to be interpreted with some care. Nevertheless, the size of the treatment effect within the first 18 months render positive effects on cumulated integration rates even in the very long-run rather unlikely. This holds true in particular due to the rather high age of participants that limits the possibilities to achieve long-run returns.

From a methodological point of view, it is worth noticing that differences between both matching estimators are negligible. Moreover, matching quality appears to be rather good, with standardized biases below 5% for almost all variables (median bias for radius matching (360 days after programme start) = 3,1). The small difference between radius matching and entropy balancing (where

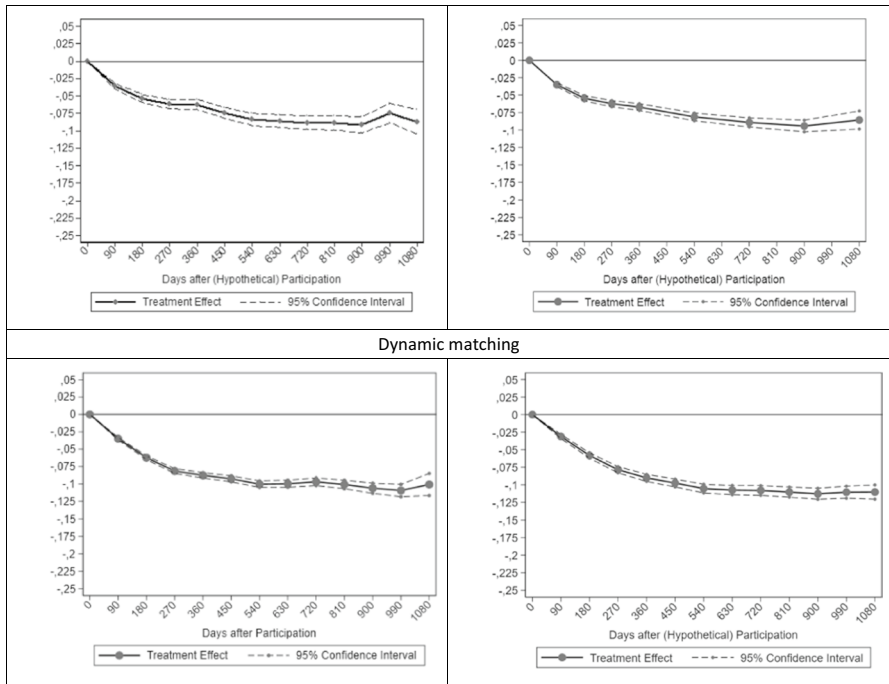


Fig. 1 ATT on labour market integration, estimated by radius matching (left panel) and entropy balancing (right panel), hypothetical start date approach. The lower part shows the results of dynamic matching. Estimated treatment effects at different points after programme start. Integration measured as having a non-subsidised job in the first labour market subject to social security. Source: Integrated Employment Biographies (V11.01.00), own calculations

standardized biases are almost zero for all variables by construction) further refute concerns that the small remaining imbalances play a major role for the estimated treatment effects. More importantly, the differences between the two approaches to defining non-treatment are very small, too. The comparatively small difference is likely due to the large share of the treatment group which enters the treatment in the first months after the end of the activation period. Therefore, the difference between the two approaches are smaller compared to settings where selection into treatment is less densely distributed and more spread out over a longer time span.

The results are fairly similar to the ones of previous research on JCS, both in terms of the direction of the effect as well as the temporal pattern. The negative effects are particularly strong at the beginning but get weaker over time, suggesting that lock-in effects play a major role. This reinforces the concern that the idea of the programme was not implemented successfully. Treated units would have had a non-negligible chance of finding a job without treatment participation, implying that the targeting mechanism was not implemented rigorously. At the same time, comparing the magnitude of the effect to the one of previous interventions reveals that it tends to be slightly weaker. As outlined in the literature

section, Hujer and Thomsen (2010) report negative employment effects of over 20% points on average for a JCS in Germany, and over 15% points for those who have entered the programme during their fifth quarter of unemployment. The impact estimates reported in our study are substantially smaller. It has to be considered that comparisons of treatment effects between programmes have to be drawn with some care, as the programmes may not only differ in their specific rules and implementation, but also in terms of the institutional setting and/or the macroeconomic conditions. As the latter could affect the effect of the programme in turn, differences between impact estimates cannot easily be ascribed to differences in the content and implementation of the programmes themselves. Nevertheless, the comparison suggests that the special institutional feature might have contributed to a weaker treatment effect. At the same time, the selection might not have been rigorous enough in order to avoid lock-in effects completely.

At the same time, it is still possible that the programme has a positive effect at least for those programme participants who are in a particularly disadvantaged position. Therefore, we repeat the analysis for five subgroups that are distinguished according to the subjective assessment of their future labour market prospects made by their counsellor at the local employment agency. Figure 2 summarizes the results starting with the group with the most optimistic assessment (upper right panel) to the least optimistic assessment (lower right corner). The upper left corner again shows the results for the whole sample for sake of comparison.

As expected, the results are even more strongly negative for participants with more optimistic assessments by the caseworkers, but close to zero or even positive (at some points in time) for the two groups with the least optimistic one. Apparently, filtering out individuals by means of the outlined pre-selection mechanism as such seems to be insufficient. However, additionally focussing on those with low integration chances based on the subjective assessment at least avoids negative effects. Once again, these results are qualitatively robust when using entropy balancing instead of radius matching (see Fig. 8). The importance of targeting is further substantiated by looking at the estimated outcome means for the control group (see Fig. 9). The control group from the two highest categories (subjective assessment: very good/good) reaches integration rates of 22% after 360 days. Considering that participants often stop looking for regular employment during the JCS, negative programme effects can rarely be avoided for groups with positive or very positive assessment. In contrast, the control group with mixed, bad and very bad assessment reaches integration rates of about 8, 3 and 2%, suggesting that lock-in effects play a minor role for those with a bad or very bad subjective assessment.

It is worth noticing that we arrive at a similar conclusion if we distinguish the sample with respect to previous employment history (e.g. having ever been employed during the last seven years, see Fig. 3). However, the effects remain negative even for individuals with long and continuous unemployment spells. Apparently, filtering out the suitable target group has to rely on a more complex procedure rather than just one straightforward criterion such as unemployment experience. This finding is again reinforced by the integration rates of the control group, which still amount to 8% for those without employment in the last seven years.

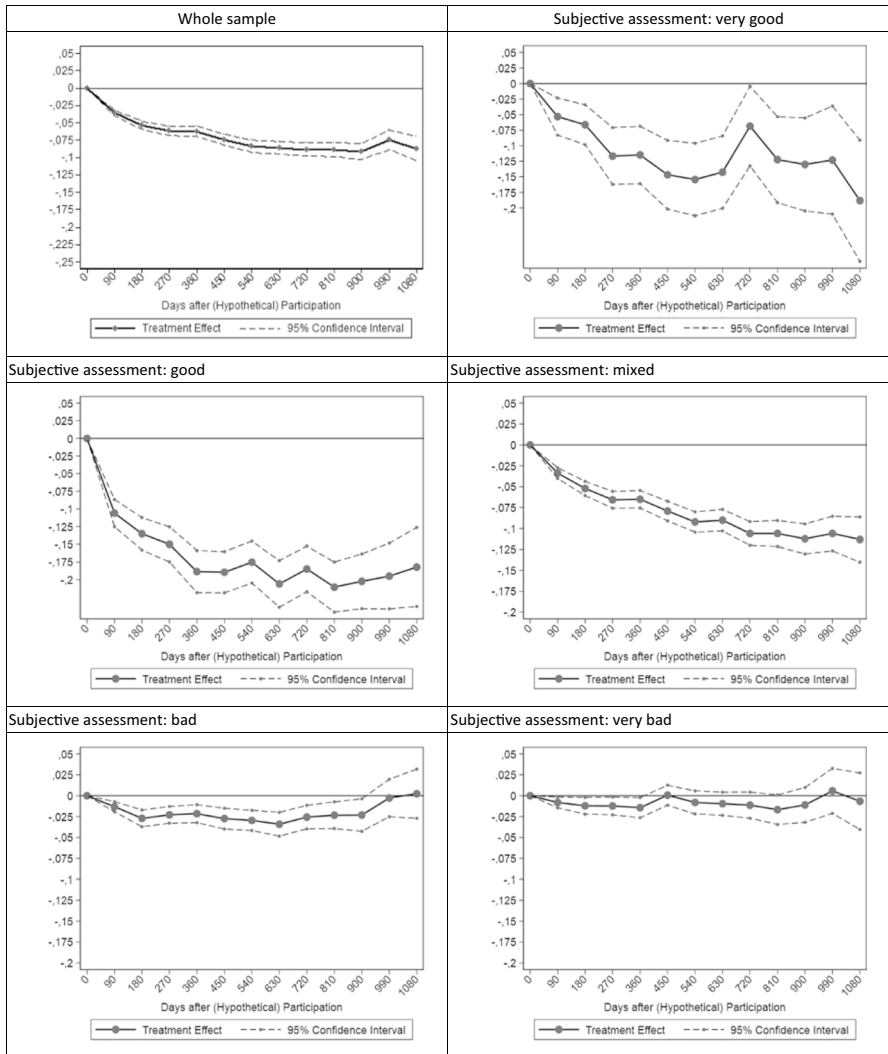


Fig. 2 ATT (entropy balancing matching) depending on assessment of the counsellor from the local employment agency. Estimated treatment effects at different points after programme start. The upper left panel shows the result for the whole sample. The remaining five panels show the results for five different assessments, from the most favourable (upper right corner) to the worst assessment. Estimated outcome means for the control group after 360 days are 0.22 (very good), 0.22 (good), 0.08 (mixed), 0.03 (bad) and 0.02 (very bad). Source: Integrated Employment Biographies (V11.01.00), own calculations

Robustness and Sensitivity Analyses

As outlined in the previous sections, the results appear to be robust with regard to different estimation techniques, but may still be subject to endogenous selection.

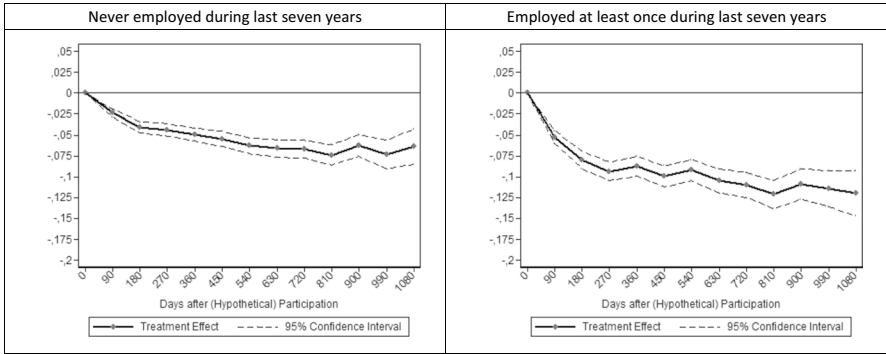


Fig. 3 ATT (radius matching) depending on past employment biography. Estimated treatment effects at different points after programme start. Integration measured as having a non-subsidised job in the first labour market subject to social security. Estimated outcome means for the control group without employment in last seven years is 0.08 (employed at least once: 0.14). Source: Integrated Employment Biographies (V11.01.00), own calculations

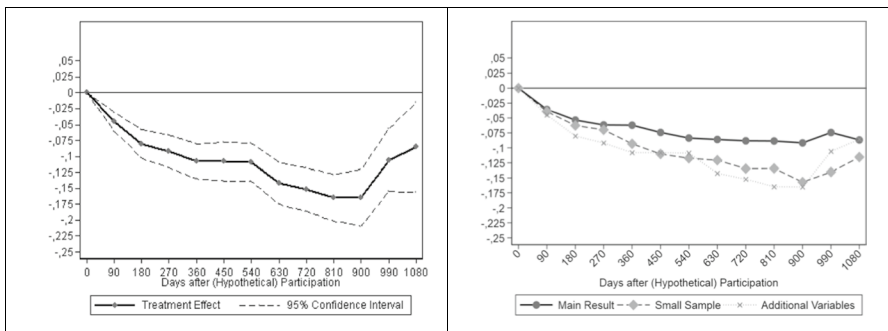


Fig. 4 ATT (radius matching) with additional variables (left panel), compared to main results and small sample without additional variables (right panel). Estimated treatment effects at different points after programme start. Integration measured as having a non-subsidised job in the first labour market subject to social security. Source: Integrated Employment Biographies (V11.01.00) and IAW employee survey, own calculations

We therefore repeat the analysis for the subsample of individuals who participated in the additional survey (see Fig. 4).

The left panel shows the results for the estimation with survey variables, the right one compares the estimation with the large sample and the survey sample without additional variables to check whether possible differences are due to sample selection of survey participation. At first glance, the new variables seem to matter as the treatment effect gets even more negative. However, comparing this estimation with the one with the survey sample without additional variables reveals that this difference is entirely due to sample selection. If the analysis is restricted to the small sample, the effects with and without the additional variables rarely differ. The underestimation of the effects in the survey sample is

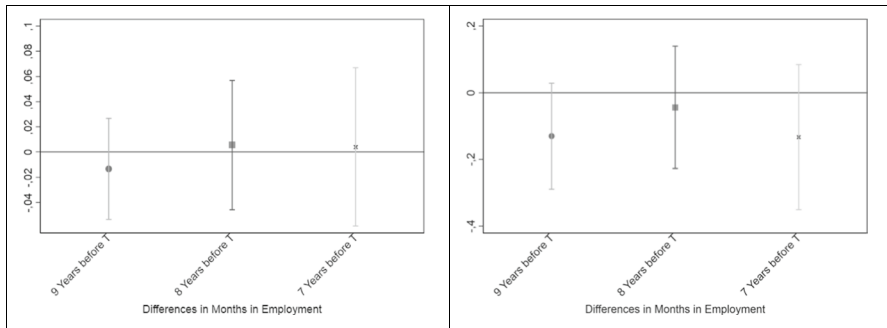


Fig. 5 Placebo test with for the whole sample (left panel) and the small sample with additional variables (right panel). Outcomes are defined as months in employment 10–9, 9–8 and 8–7 years before the start of the treatment. Estimated treatment effects at different points after programme start. Integration measured as having a non-subsidised job in the first labour market subject to social security. Source: Integrated Employment Biographies (V11.01.00) and IAW employee survey, own calculations

due to an overrepresentation of individuals with higher labour market attachment (shorter cumulated unemployment in the past, higher education and qualification, more favourable subjective assessment of future labour market chances), for who the negative treatment effect is stronger. In this regard, it can be concluded that conditioning on further covariates does not change the results. From a methodological point of view, this finding implies that endogenous selection may be weaker than expected and substantiates the credibility of the CIA. This complements the findings of Caliendo et al. (2017) who confirm that additional (usually unobservable) variables on psycho-social characteristics do not matter if high quality register data are available, even though these additional variables appear to predict programme participation (propensity score estimation with survey variables see Table 3). It is worth noticing that Caliendo et al. (2017) do report some differences for wage subsidies where employers are involved in the selection process. While employer selection plays a role in our institutional setting, too, we may arrive at different results here since JCS employers may have a different utility function and might therefore be less selective. Moreover, the problem of endogenous selection might be weakened by exogenous factors that determine participation in the JCS. Since the number of observations with survey information is considerably lower, it is worth noting that differences between matching estimators remain limited and both algorithms still achieve sufficient matching quality (Fig. 9).

To substantiate the credibility of our research design even further, we have conducted a series of placebo tests. To this end, we define the number of months in employment in the time span between eight and seven, nine and eight and ten and nine years before programme start as outcomes. The employment biography used as covariates lasts until seven years before treatment start, i.e. the placebo outcomes are not used as matching variables. As Fig. 5 shows, the placebo effects are close to zero and significant for all three placebo outcomes, regardless of whether survey variables are included or not. In this regard, the sensitivity test with additional variables

and the placebo test are consistent in reinforcing the credibility of the CIA in our setting (Table 4).

Moreover, we have implemented two recent estimation techniques that may be suitable when the conditional independence assumption is violated but external instruments are not available. These methods have been developed and implemented earlier in Chen and Wang (2020), Millimet and Tchernis (2013) as well as McCarthy et al. (2014). The first method (referred to as minimum-biased estimator by Millimet and Tchernis 2013) restricts the sample to observations close to the bias-minimizing propensity score. It hereby reduces the number of observations (loss of efficiency) but also reduces possible biases. The bias-corrected approach extends the minimum-biased approach further by introducing a bias-correction procedure. In a nutshell, it constructs internal instruments by exploiting heteroskedasticity of the error term for identification. The results (see Table 5) point in the same direction, the treatment effect remains strongly negative (point estimates are much larger, but inflated point estimates are common within this method (see McCarthy et al. 2014 for an in-depth discussion and empirical examples). These results confirm that the negative treatment effects are not driven by negative unobserved selection into treatment.

Summary, Discussion and Conclusion

This paper has assessed the impact of a recent, large-scale, and innovative JCS in Germany. Employing a special selection mechanism, the JCS under discussion aimed at overcoming the shortcomings of previous JCS. In particular, targeting very hard to place individuals was supposed to avoid lock-in effects, as integration rates of the target group would have been very low in case of non-participation. Despite this innovative approach, the average treatment effect appears to be strongly negative. We argue that this is partly due to the involvement of JCS employers in picking participants, as employers are likely to choose participants with more favourable characteristics. At the same time, distinguishing with respect to the assessment of future labour market chances as estimated by the caseworkers in the local employment agencies shows that the treatment effect is at least not negative for job seekers with particularly low labour market attachment. These findings are robust to the use of different matching algorithms. Moreover, placebo tests as well as the inclusion of usually unobservable survey variables refute concerns about endogenous selection.

Thinking about implications for future research and policy-making, the results reveal an ambiguous picture. On the one hand, it has to be admitted that targeting JCS via employment history alone or the outlined pre-selection mechanism appears to be insufficient. Other labour market programmes such as training programmes or subsidized employment, which have been shown to foster labour market integration of certain groups (Brown and Koettl 2015; Bellmann et al. 2018; Ahmad et al. 2019), do not seem to be suitable for job seekers with very low labour market attachment. As a result, supporting the reintegration of very hard to place workers in the labour market remains an unsettled issue. On the other hand, our sub-group analysis reveals that JCS may still be an option if the targeting strategy succeeds to identify job seekers with very little employment chances. Doing so, it appears to be

necessary to reinforce the targeting at all stages of the selection process. While the explicit involvement of JCS-employers (as in our case) may be a special scenario at first glance, cream-skimming may be a relevant problem in other contexts, too. To ensure effective targeting, future JCS could either avoid the involvement of employers or create incentives to choose hard-to-place workers. It has to be left to future research to what extent such approaches could be successful. At the methodological level, our results complement previous research that has reinforced the credibility of the CIA if high quality register data are available. While experimental research in labour economics is growing, this suggests that observational studies can still be a reliable source for policy advice. Considering that random assignment may not be an option or not feasible in some cases, this suggests that further developing and employing non-experimental methodology should not be regarded as a dead end.

Appendix

See Figs. 6, 7, 8, 9 and 10 and Tables 1, 2, 3, 4, and 5..

Table 1 Summary statistics for outcomes and covariates of treatment and control group

Variable	Treatment group					Control group					Sig-nificance of difference
	Mean	SD	min	max	N	mean	SD	min	max	N	
Participant	1	0	1	1	12,207	0	0	0	0	51,536	
Education: low	0.2	0.4	0	1	12,207	0.26	0.44	0	1	51,535	**
Education: high school degree	0.67	0.47	0	1	12,207	0.6	0.49	0	1	51,535	**
Education: college degree	0.05	0.22	0	1	12,207	0.04	0.19	0	1	51,535	**
Age group: 25–29	0.08	0.27	0	1	12,207	0.19	0.4	0	1	51,536	**
Age group: 30 to 34	0.1	0.3	0	1	12,207	0.14	0.35	0	1	51,536	**
Age group: 40 to 44	0.19	0.39	0	1	12,207	0.17	0.38	0	1	51,536	**
Age group: 45 to 49	0.26	0.44	0	1	12,207	0.19	0.39	0	1	51,536	*
Age group: 50 to 57	0.2	0.4	0	1	12,207	0.14	0.34	0	1	51,536	**
Age group: 58 and older	0.05	0.22	0	1	12,207	0.03	0.17	0	1	51,536	**
Married	0.33	0.47	0	1	12,207	0.33	0.47	0	1	51,536	
Divorced/widowed/separated	0.25	0.43	0	1	12,207	0.22	0.41	0	1	51,536	**
Migration background	0.05	0.23	0	1	12,207	0.11	0.31	0	1	51,536	**
Health problems	0.26	0.44	0	1	12,207	0.24	0.43	0	1	51,528	**
Blue collar worker	0.02	0.14	0	1	12,207	0.24	0.43	0	1	51,536	**

Table 1 (continued)

Variable	Treatment group					Control group					Sig- nificance of difference
	Mean	SD	min	max	N	mean	SD	min	max	N	
Professional edu- cation: none	0.49	0.5	0	1	12,207	0.53	0.5	0	1	51,536	**
Required educa- tion: missing	0.34	0.47	0	1	12,207	0.31	0.46	0	1	51,536	**
Complexity of last job: basic	0.69	0.46	0	1	12,207	0.68	0.46	0	1	51,536	
Complexity of last job: complex	0.08	0.27	0	1	12,207	0.07	0.26	0	1	51,536	*
Complexity of last job: highly complex	0.1	0.3	0	1	12,207	0.1	0.3	0	1	51,536	
Sector of the industry: production	0.44	0.5	0	1	12,207	0.41	0.49	0	1	51,536	*
Qualification match	0.05	0.22	0	1	12,207	0.05	0.22	0	1	51,536	
daily wage of last job	23.67	11.86	0	371.59	12,201	23.65	17.24	0	1881.6	49,699	
Special status	0.11	0.31	0	1	12,207	0.16	0.37	0	1	51,536	**
Relaxed ben- efit entitlement condition (1)	0.01	0.09	0	1	12,207	0.02	0.13	0	1	51,536	**
Relaxed ben- efit entitlement condition (2)	0.06	0.23	0	1	12,207	0.05	0.21	0	1	51,536	**
Previously terminated programme participation	0.03	0.18	0	1	12,207	0.05	0.22	0	1	51,536	**
Previously completed programme successfully	0.71	0.46	0	1	12,207	0.63	0.48	0	1	51,536	**
Result of ALMP: missing	0.25	0.43	0	1	12,207	0.31	0.46	0	1	51,536	**
Has child below age 3	0.02	0.12	0	1	12,207	0.04	0.19	0	1	51,536	**
Has child aged 3 to 6	0.03	0.16	0	1	12,207	0.04	0.2	0	1	51,536	**
Has child aged 6 to 10	0.06	0.24	0	1	12,207	0.08	0.27	0	1	51,536	**
Has child aged 10 to 15	0.09	0.28	0	1	12,207	0.1	0.3	0	1	51,536	**
N persons in household: 2	0.25	0.43	0	1	12,207	0.22	0.41	0	1	51,536	**
N persons in household: 3	0.13	0.34	0	1	12,207	0.14	0.35	0	1	51,536	*

Table 1 (continued)

Variable	Treatment group					Control group					Sig-nificance of difference
	Mean	SD	min	max	N	mean	SD	min	max	N	
N persons in household: 4	0.08	0.26	0	1	12,207	0.09	0.29	0	1	51,536	**
N persons in household: 5	0.04	0.2	0	1	12,207	0.06	0.24	0	1	51,536	**
Lone parent	0.17	0.38	0	1	12,207	0.17	0.38	0	1	51,536	
Subjective assessment: Very good	0.03	0.16	0	1	12,207	0.08	0.27	0	1	51,536	**
Subjective assessment: mixed	0.48	0.5	0	1	12,207	0.32	0.47	0	1	51,536	**
Subjective assessment: bad	0.21	0.41	0	1	12,207	0.17	0.38	0	1	51,536	**
Subjective assessment: very bad hard to place	0.15	0.36	0	1	12,207	0.15	0.36	0	1	51,536	
Subjective assessment: missing	0.03	0.16	0	1	12,207	0.1	0.3	0	1	51,536	**
Common administration municipality/employment agency	0.04	0.19	0	1	12,207	0.06	0.23	0	1	51,536	**
Region of residence: Bavaria	0.07	0.26	0	1	12,207	0.13	0.34	0	1	51,536	**
Region of residence: Brandenburg	0.09	0.28	0	1	12,207	0.06	0.24	0	1	51,536	**
Region of residence: Hessen	0.03	0.17	0	1	12,207	0.04	0.21	0	1	51,536	**
Region of residence: Lower Saxony/Bremen	0.07	0.25	0	1	12,207	0.07	0.26	0	1	51,536	
Region of residence: North	0.12	0.32	0	1	12,207	0.09	0.28	0	1	51,536	**
Region of residence: Northrhine-Westfalia	0.09	0.29	0	1	12,207	0.09	0.29	0	1	51,536	
Region of residence: Rhineland-Palatinate/Saarland	0.1	0.3	0	1	12,207	0.07	0.25	0	1	51,536	*
Region of residence: Saxony	0.1	0.3	0	1	12,207	0.06	0.25	0	1	51,536	**

Table 1 (continued)

Variable	Treatment group					Control group					Sig- nificance of difference
	Mean	SD	min	max	N	mean	SD	min	max	N	
Region of residence: Lower-Saxony-Anhalt/Thuringia	0.29	0.46	0	1	12,207	0.32	0.47	0	1	51,536	**
ESF special region: yes	0.57	0.5	0	1	12,207	0.5	0.5	0	1	51,536	**
Months employed 1 year before program start	0.35	1.56	0	12	12,207	0.36	1.48	0	12	51,536	
Months employed 2–4 years before program start	1.9	5.29	0	36	12,207	2.68	6.55	0	36	51,536	**
Months employed 5–7 years before program start	3.91	8.19	0	36	12,207	4.18	8.59	0	36	51,536	**
Months unemployed 1 year before program start	9.24	3.74	0	12	12,207	9.53	3.62	0	12	51,536	**
Months unemployed 2–4 years before program start	26.73	9.56	0	36	12,207	25.01	11.27	0	36	51,536	**
Months unemployed 5–7 years before program start	16.25	8.12	0	36	12,207	14.61	8.98	0	36	51,536	**
Months in ALMP program 1 year before program start	0.45	1.53	0	12	12,207	0.55	1.77	0	12	51,536	**
Months in ALMP program 2–4 year before program start	1.6	3.92	0	36	12,207	1.94	4.85	0	36	51,536	**
Months in ALMP program 5–7 year before program start	8.35	8.02	0	36	12,207	6.43	7.66	0	36	51,536	**
Ever been employed in last seven years	0.43	0.49	0	1	12,207	0.45	0.5	0	1	51,536	**
Regional UE rate	10.5	3.3	3.1	17.4	11,985	10.0	3.3	3.1	17.4	50,603	**

Variable	Treatment group					Control group					Sig-nificance of difference
	Mean	SD	min	max	N	mean	SD	min	max	N	
Regional employment rate	50.5	3.6	38.8	59.1	11,985	50.3	3.6	38.5	59.1	50,603	**
Regional Population density (in thousands)	0.73	0.76	0.04	4.4	11,933	0.81	0.78	0.04	4.4	49,463	**
Regional GDP(in thousands)	53.8	7.3	42.7	79.0	12,093	54.2	7.9	42.7	79.0	50,922	**
Unemployment duration before program start	55.3	26.52	6	111	12,042	49.94	26.93	6	114	50,461	**

Stars indicate level of significance (* $p < 0.05$; ** $p < 0.01$)

Source: Integrated Employment biographies (V11.01.00); own calculations

Table 2 Summary statistics for outcomes and covariates of treatment and control group, survey variables

Variable	Treatment group					Control group					Significance of difference
	Mean	SD	min	max	N	mean	SD	min	max	N	
Subjective health	2.44	1.16	1	5	1308	2.56	1.27	1	5	3658	**
Subjective resilience	3.45	0.72	1	4	1308	3.35	0.85	1	4	3658	**
Self-assessed reading skills	1.92	0.87	1	6	1308	2.04	0.97	1	6	3658	**
Self-assessed maths skills	2.33	0.96	1	6	1308	2.38	0.98	1	6	3658	
Self-assessed internet skills	2.69	1.55	1	6	1308	2.68	1.55	1	6	3658	
Driver's license	0.59	0.49	0	1	1308	0.56	0.5	0	1	3658	
Peer group: many social assistance recipients	0.54	0.5	0	1	1308	0.53	0.5	0	1	3658	
Peer group: successful in professional life	0.62	0.49	0	1	1308	0.64	0.48	0	1	3658	
Peer group: interested in own professional life	0.85	0.36	0	1	1308	0.83	0.37	0	1	3658	
Feel support by partner	0.3	0.46	0	1	1308	0.33	0.47	0	1	3658	**
Feel support by family	0.49	0.5	0	1	1308	0.49	0.5	0	1	3658	
Feel support by friends	0.4	0.49	0	1	1308	0.4	0.49	0	1	3658	
Personality traits (component 1)	-0.1	1.49	-2.4	5.3	1308	-0.02	1.6	-2.4	9.6	3658	
Personality traits (component 2)	0.01	1.14	-5.1	4	1308	-0.05	1.12	-5.1	4.6	3658	
Support (component 1)	0.02	1	-3	8.5	1308	0.02	1	-3	10.4	3658	
Support (component 2)	-0.1	1.13	-14.7	6.1	1308	-0.03	0.97	-8.9	7.9	3658	**
Personal problems (component 1)	-0.06	1.24	-0.8	7.3	1308	0.06	1.35	-0.8	9	3658	**
Personal problems (component 2)	0.05	1.03	-3.3	6.4	1308	-0.01	1.05	-3.3	6.2	3658	
Number of job interviews last six months	0.64	0.85	0	3.91	1308	0.71	0.88	0	3.91	3658	*
Concession: commuting time > 1.5 h	2.66	1.1	1	4	1308	2.67	1.13	1	4	3658	
Concession: shift work	2.19	1.2	1	4	1308	2.2	1.21	1	4	3658	
Concession: unpleasant work environment	1.94	1.06	1	4	1308	1.92	1.05	1	4	3658	

Table 2 (continued)

Variable	Treatment group					Control group					Significance of difference
	Mean	SD	min	max	N	mean	SD	min	max	N	
Concession: move to other city	3.24	1.06	1	4	1308	3.16	1.09	1	4	3658	*
Concession: hourly wage < 8€	2.71	1.1	1	4	1308	2.71	1.12	1	4	3658	
Feeling of social participation	6.27	2.64	0	10	1308	6.33	2.64	0	10	3658	
Coping with current situation	5.75	2.38	0	10	1308	5.92	2.47	0	10	3658	

Stars indicate level of significance (* $p < 0.05$; ** $p < 0.01$). Source: Integrated Employment biographies (V11.01.00) and IAW employee survey; own calculations

Table 3 Propensity score estimation at different points in time after treatment start

	(180 days)	(360 days)	(720 days)	(1080 days)
Level of education: 2	– 0.0133 (0.0265)	– 0.0162 (0.0267)	– 0.0229 (0.0274)	– 0.0403 (0.0398)
Level of education: 3	0.1226*** (0.0255)	0.1200*** (0.0256)	0.1172*** (0.0263)	0.1227** (0.0379)
Level of education: 5	0.3565*** (0.0440)	0.3545*** (0.0442)	0.3440*** (0.0453)	0.3506*** (0.0655)
Age group: 25–29	– 0.3066*** (0.0298)	– 0.3039*** (0.0300)	– 0.3283*** (0.0309)	– 0.3675*** (0.0451)
Age group: 30 to 34	– 0.0916*** (0.0272)	– 0.0902*** (0.0274)	– 0.1043*** (0.0282)	– 0.1346** (0.0414)
Age group: 40 to 44	0.0921*** (0.0241)	0.0915*** (0.0243)	0.0829*** (0.0249)	0.1272*** (0.0361)
Age group: 45 to 49	0.1701*** (0.0239)	0.1726*** (0.0240)	0.1724*** (0.0246)	0.2035*** (0.0356)
Age group: 50 to 57	0.1974*** (0.0266)	0.2006*** (0.0267)	0.2098*** (0.0274)	0.2612*** (0.0393)
Age group: 58 and older	0.3312*** (0.0414)	0.3414*** (0.0415)	0.3612*** (0.0426)	0.4111*** (0.0594)
Married	– 0.0857*** (0.0245)	– 0.0914*** (0.0246)	– 0.0874*** (0.0254)	– 0.0726* (0.0361)
Divorced/widowed/separated	– 0.0761*** (0.0189)	– 0.0786*** (0.0190)	– 0.0754*** (0.0195)	– 0.0635* (0.0281)
Migration background	– 0.2830*** (0.0270)	– 0.2833*** (0.0271)	– 0.2802*** (0.0278)	– 0.3481*** (0.0407)
Health problems	– 0.0187 (0.0160)	– 0.0173 (0.0161)	– 0.0097 (0.0165)	0.0235 (0.0235)
Blue collar worker	– 1.6344*** (0.0301)	– 1.6313*** (0.0302)	– 1.6142*** (0.0308)	– 1.4915*** (0.0408)
Professional education: none	0.0280 (0.0200)	0.0285 (0.0201)	0.0284 (0.0206)	0.0392 (0.0293)
Required education: missing	0.0101 (0.0220)	0.0094 (0.0221)	0.0128 (0.0227)	0.0371 (0.0321)
Complexity of last job: basic	0.0067 (0.0205)	0.0114 (0.0206)	0.0105 (0.0212)	– 0.0431 (0.0303)
Complexity of last job: complex	0.0762* (0.0314)	0.0838** (0.0315)	0.0780* (0.0325)	0.0187 (0.0470)
Complexity of last job: highly complex	0.0957** (0.0309)	0.1000** (0.0311)	0.1073*** (0.0319)	0.0735 (0.0458)
Sector of the industry: production	0.0743*** (0.0144)	0.0723*** (0.0145)	0.0680*** (0.0149)	0.0636** (0.0213)
Qualification match	0.0340 (0.0408)	0.0378 (0.0410)	0.0296 (0.0424)	0.0772 (0.0600)

Table 3 (continued)

	(180 days)	(360 days)	(720 days)	(1080 days)
Daily wage	0.0052*** (0.0005)	0.0054*** (0.0005)	0.0050*** (0.0005)	0.0043*** (0.0008)
Special status	- 0.1196*** (0.0333)	- 0.1154*** (0.0334)	- 0.1094** (0.0343)	- 0.1444** (0.0497)
Relaxed benefit entitlement condition (1)	0.1697* (0.0733)	0.1569* (0.0738)	0.1273+ (0.0763)	0.1556 (0.1129)
Relaxed benefit entitlement condition (2)	0.1419** (0.0512)	0.1417** (0.0515)	0.1384** (0.0530)	0.2011** (0.0759)
Previously terminated programme participation	- 0.0274 (0.0805)	- 0.0237 (0.0811)	- 0.0438 (0.0844)	- 0.1784 (0.1244)
Previously completed programme successfully	0.0945 (0.0741)	0.0948 (0.0746)	0.0901 (0.0777)	0.0351 (0.1138)
Previous programme result: not known	0.0178 (0.0749)	0.0174 (0.0755)	0.0103 (0.0786)	- 0.0558 (0.1150)
Has child below 3	- 0.1890*** (0.0468)	- 0.1962*** (0.0472)	- 0.1850*** (0.0490)	- 0.1163 (0.0710)
Has child below age 3 to 6	0.0435 (0.0398)	0.0301 (0.0402)	0.0135 (0.0418)	- 0.0229 (0.0621)
Has child aged 6 to 10	0.0030 (0.0294)	- 0.0010 (0.0296)	- 0.0133 (0.0307)	0.0091 (0.0445)
Has child aged 10 to 15	- 0.0019 (0.0264)	- 0.0080 (0.0266)	- 0.0215 (0.0276)	- 0.0654 (0.0400)
N persons in household: 2	0.0169 (0.0246)	0.0227 (0.0247)	0.0222 (0.0255)	0.0238 (0.0365)
N persons in household: 3	0.0006 (0.0282)	0.0083 (0.0284)	0.0087 (0.0292)	0.0517 (0.0416)
N persons in household: 4	- 0.0023 (0.0340)	0.0057 (0.0342)	0.0014 (0.0353)	0.0274 (0.0503)
N persons in household: 5 or more	- 0.0932* (0.0409)	- 0.0842* (0.0411)	- 0.0793+ (0.0423)	0.0147 (0.0600)
Lone parent	- 0.0248 (0.0256)	- 0.0270 (0.0258)	- 0.0124 (0.0266)	- 0.0142 (0.0381)
Subjective assessment: Very good	- 0.3713*** (0.0365)	- 0.3693*** (0.0366)	- 0.3706*** (0.0375)	- 0.3738*** (0.0551)
Subjective assessment: Mixed	0.4139*** (0.0217)	0.4110*** (0.0218)	0.4063*** (0.0223)	0.4746*** (0.0330)
Subjective assessment: Bad	0.3919*** (0.0251)	0.3885*** (0.0252)	0.3759*** (0.0259)	0.4150*** (0.0382)
Subjective assessment: Very bad	0.3661*** (0.0270)	0.3588*** (0.0271)	0.3517*** (0.0278)	0.4015*** (0.0404)
Subjective assessment: not known	- 0.3742*** (0.0371)	- 0.3759*** (0.0372)	- 0.3778*** (0.0379)	- 0.3884*** (0.0542)

Table 3 (continued)

	(180 days)	(360 days)	(720 days)	(1080 days)
Common administration	– 0.0288	– 0.0339	– 0.0540	– 0.0362
municipality/employment agency	(0.0406)	(0.0408)	(0.0420)	(0.0571)
region of residence: Bavaria	– 0.0955*	– 0.0978*	– 0.1212**	– 0.1382*
	(0.0401)	(0.0404)	(0.0415)	(0.0616)
region of residence: Brandenburg	0.2009***	0.2051***	0.2296***	0.2616**
	(0.0564)	(0.0567)	(0.0583)	(0.0882)
region of residence: Hessen	0.0106	0.0146	0.0184	0.0854
	(0.0472)	(0.0475)	(0.0486)	(0.0694)
region of residence: Lower Saxony/ Bremen	0.1445***	0.1465***	0.1561***	0.1151+
	(0.0426)	(0.0428)	(0.0439)	(0.0666)
region of residence: North	0.2816***	0.2838***	0.2827***	0.1988*
	(0.0479)	(0.0481)	(0.0495)	(0.0780)
region of residence: Northrhine-Westfalia	0.1706***	0.1745***	0.1828***	0.2058**
	(0.0424)	(0.0426)	(0.0437)	(0.0647)
region of residence: Rhineland-Palatinate/ Saarland	0.3844***	0.3885***	0.3975***	0.5536***
	(0.0410)	(0.0412)	(0.0423)	(0.0632)
region of residence: Saxony	0.2672***	0.2691***	0.2868***	0.2155*
	(0.0580)	(0.0582)	(0.0599)	(0.0921)
region of residence: Lower-Saxony- Anhalt/Thuringia	0.0387	0.0391	0.0491	0.0674
	(0.0546)	(0.0549)	(0.0565)	(0.0864)
ESF special region: yes	– 0.1641***	– 0.1636***	– 0.1714***	– 0.2608***
	(0.0426)	(0.0428)	(0.0443)	(0.0716)
Months employed 1 year before program start	– 0.0018	– 0.0023	– 0.0057	– 0.0093
	(0.0055)	(0.0055)	(0.0057)	(0.0087)
Months employed 2–4 years before program start	– 0.0031+	– 0.0032+	– 0.0047**	– 0.0062*
	(0.0017)	(0.0017)	(0.0017)	(0.0025)
Months employed 5–7 years before program start	0.0045***	0.0045***	0.0054***	0.0059**
	(0.0012)	(0.0012)	(0.0013)	(0.0018)
Months unemployed 1 year before program start	– 0.0320***	– 0.0324***	– 0.0333***	– 0.0333***
	(0.0024)	(0.0024)	(0.0025)	(0.0036)
Months employed 2–4 years before program start	– 0.0016	– 0.0016	– 0.0004	0.0015
	(0.0010)	(0.0010)	(0.0011)	(0.0015)
Months employed 5–7 years before program start	0.0042***	0.0042***	0.0049***	0.0064***
	(0.0011)	(0.0011)	(0.0011)	(0.0016)
Months in program participation 1 year before program start	– 0.0294***	– 0.0305***	– 0.0295***	– 0.0338***
	(0.0047)	(0.0048)	(0.0049)	(0.0071)
Months in program participation 2–4 years before program start	– 0.0037*	– 0.0038*	– 0.0037+	– 0.0015
	(0.0018)	(0.0018)	(0.0019)	(0.0027)
Months in program participation 5–7 years before program start	0.0095***	0.0096***	0.0100***	0.0105***
	(0.0011)	(0.0011)	(0.0011)	(0.0016)
Ever been employed in last 7 years	0.0139	0.0161	0.0020	– 0.0034
	(0.0184)	(0.0185)	(0.0190)	(0.0273)

Table 3 (continued)

	(180 days)	(360 days)	(720 days)	(1080 days)
Regional UE rate	0.0388*** (0.0040)	0.0387*** (0.0041)	0.0382*** (0.0042)	0.0639*** (0.0062)
Regional employment rate	0.0327*** (0.0032)	0.0331*** (0.0033)	0.0334*** (0.0034)	0.0379*** (0.0050)
Regional Population density	− 0.0000* (0.0000)	− 0.0000+ (0.0000)	− 0.0000 (0.0000)	− 0.0001*** (0.0000)
Regional GDP	0.0000*** (0.0000)	0.0000*** (0.0000)	0.0000*** (0.0000)	0.0000 (0.0000)
UE duration before treatment start	0.0017*** (0.0003)	0.0017*** (0.0003)	0.0002 (0.0003)	− 0.0030*** (0.0004)
N	51,394	51,010	48,152	23,763
chi2	9055.78	9010.42	8561.51	4608.10
r2_p	0.16	0.17	0.17	0.18
aic	45,982.22	45,591.98	43,221.82	20,938.91

Results from probit estimations. Standard errors in parentheses. + $p < 0.10$. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$

Table 4 Propensity score estimation with survey and register variables variable

	(180)	(360)	(720)	(1080)
Subjective health	– 0.0334 (0.0219)	– 0.0346 (0.0220)	– 0.0312 (0.0228)	– 0.0730 ⁺ (0.0401)
Subjective resilience	0.1317 ^{***} (0.0336)	0.1308 ^{***} (0.0338)	0.1302 ^{***} (0.0353)	0.0730 (0.0621)
Self-assessed reading skills	– 0.0809 ^{**} (0.0306)	– 0.0796 ^{**} (0.0307)	– 0.0813 [*] (0.0317)	– 0.0695 (0.0534)
Self-assessed math skills	0.0400 (0.0277)	0.0385 (0.0278)	0.0272 (0.0290)	0.0874 ⁺ (0.0488)
Self-assessed internet skills	– 0.0322 ⁺ (0.0179)	– 0.0342 ⁺ (0.0180)	– 0.0312 ⁺ (0.0187)	– 0.0713 [*] (0.0318)
Driver's license	0.0674 (0.0505)	0.0594 (0.0507)	0.0439 (0.0525)	0.1042 (0.0905)
Peer group: many social assistance recipients	– 0.1036 [*] (0.0521)	– 0.1051 [*] (0.0524)	– 0.0913 ⁺ (0.0541)	– 0.0274 (0.0927)
Peer group: successful in professional life	– 0.1083 [*] (0.0531)	– 0.0984 ⁺ (0.0534)	– 0.1002 ⁺ (0.0554)	– 0.1199 (0.0953)
Peer group: interested in own professional life	0.0510 (0.0685)	0.0517 (0.0691)	0.0263 (0.0723)	– 0.1782 (0.1252)
Feel support by partner	– 0.0401 (0.0560)	– 0.0280 (0.0562)	– 0.0471 (0.0582)	– 0.0472 (0.0982)
Feel support by family	0.0665 (0.0494)	0.0620 (0.0497)	0.0659 (0.0515)	0.2150 [*] (0.0884)
Feel support by friends	0.0002 (0.0494)	0.0039 (0.0497)	0.0043 (0.0514)	– 0.0257 (0.0870)
Personality traits (component 1)	– 0.0031 (0.0168)	– 0.0014 (0.0169)	– 0.0031 (0.0174)	– 0.0279 (0.0308)
Personality traits (component 2)	– 0.0199 (0.0232)	– 0.0198 (0.0233)	– 0.0207 (0.0241)	– 0.0296 (0.0413)
Institutional support (component 1)	0.0133 (0.0230)	0.0139 (0.0231)	0.0149 (0.0238)	0.0426 (0.0429)
Institutional support (component 2)	– 0.0406 ⁺ (0.0226)	– 0.0464 [*] (0.0232)	– 0.0478 [*] (0.0236)	– 0.1334 ^{**} (0.0488)
Personal problems (component 1)	– 0.0184 (0.0192)	– 0.0175 (0.0193)	– 0.0259 (0.0202)	– 0.0526 (0.0360)
Personal problems (component 2)	0.0411 ⁺ (0.0231)	0.0420 ⁺ (0.0232)	0.0487 [*] (0.0241)	0.0747 ⁺ (0.0425)
Number of job interviews (last 6 months)	– 0.0145 (0.0274)	– 0.0131 (0.0275)	– 0.0042 (0.0285)	0.0225 (0.0508)
Concession: commuting time > 1.5 h	0.0059 (0.0223)	0.0053 (0.0225)	0.0105 (0.0233)	0.0220 (0.0401)
Concession: Shift work	0.0052 (0.0223)	0.0030 (0.0224)	0.0083 (0.0233)	0.0474 (0.0406)

Table 4 (continued)

	(180)	(360)	(720)	(1080)
Concession: unpleasant work environment	0.0113 (0.0238)	0.0085 (0.0240)	0.0103 (0.0248)	0.0318 (0.0420)
Concession: move to different city	0.0044 (0.0232)	0.0040 (0.0233)	0.0046 (0.0242)	0.0313 (0.0426)
Concession: hourly wage > 8€	0.0033 (0.0217)	0.0043 (0.0218)	0.0105 (0.0227)	0.0274 (0.0389)
Feeling of social participation	- 0.0024 (0.0101)	- 0.0026 (0.0101)	- 0.0080 (0.0105)	- 0.0115 (0.0176)
Coping with current situation	- 0.0060 (0.0106)	- 0.0047 (0.0107)	- 0.0039 (0.0111)	0.0043 (0.0188)
N	4069	4026	3769	1383
chi2	888.92	880.51	854.11	476.18
r2_p	0.18	0.18	0.18	0.26

Propensity score estimation with register and survey variables (register variables as in baseline specification). Standard errors in parentheses. + $p < 0.10$. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$

Table 5 Summary of average treatment effects on the treated (point estimates) comparing OLS, Inverse probability weighting, and two estimators developed by Millimet and Tchernis (2013)

	OLS	IPW	MB	MB-BC
180 days	- 0.061 [- 0.064; - 0.058]	- 0.055 [- 0.058; - 0.052]	- 0.042 [- 0.046; - 0.037]	- 0.344 [- 0.371; - 0.302]
360 days	- 0.075 [- 0.079; - 0.072]	- 0.068 [- 0.072; - 0.064]	- 0.051 [- 0.057; - 0.045]	- 0.471 [- 0.505; - 0.426]

Estimation has been implemented using the `bmt` command of STATA. For the MB and MB-BC method, the chosen value for theta is 0.05, the value for P^* is equal to 0.500

Source: Integrated employment biographies (V11.01.00)

Fig. 6 Distribution of time between the end of the activation period and the start of the JCS, used as basis for the hypothetical start date approach. Source: IEB employment biographies (V11.01.00), own calculations

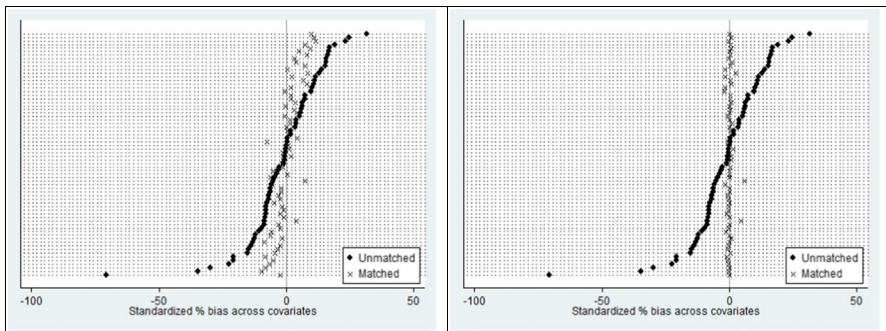
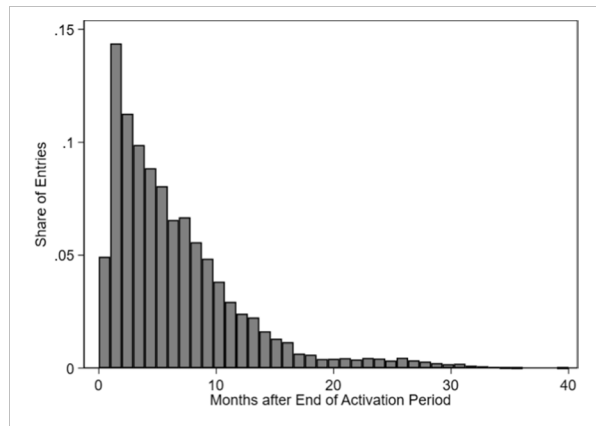


Fig. 7 Standardized bias before and after matching for radius matching (left panel) and entropy balancing (right panel) for the estimated effects 360 days after programme start). Source: IEB employment biographies (V11.01.00), own calculations

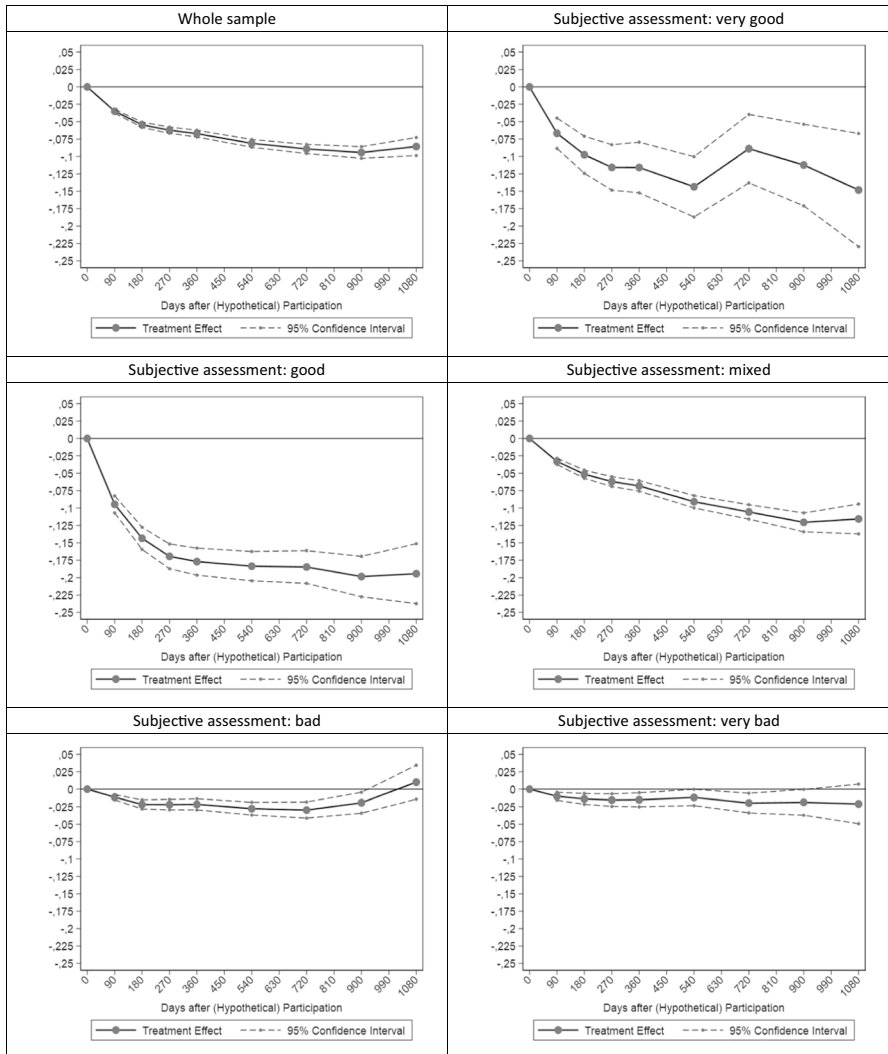


Fig. 8 ATT (radius matching) depending on assessment of the counsellor from the local employment agency. Estimated treatment effects at different points after programme start. Integration measured as having a non-subsidised job in the first labour market subject to social security. The upper left panel shows the result for the whole sample. The remaining five panels show the results for five different assessments, one (upper right corner) representing the most favourable assessment, and five (lower right corner) the least optimistic one. Source: Integrated Employment Biographies (V11.01.00), own calculations

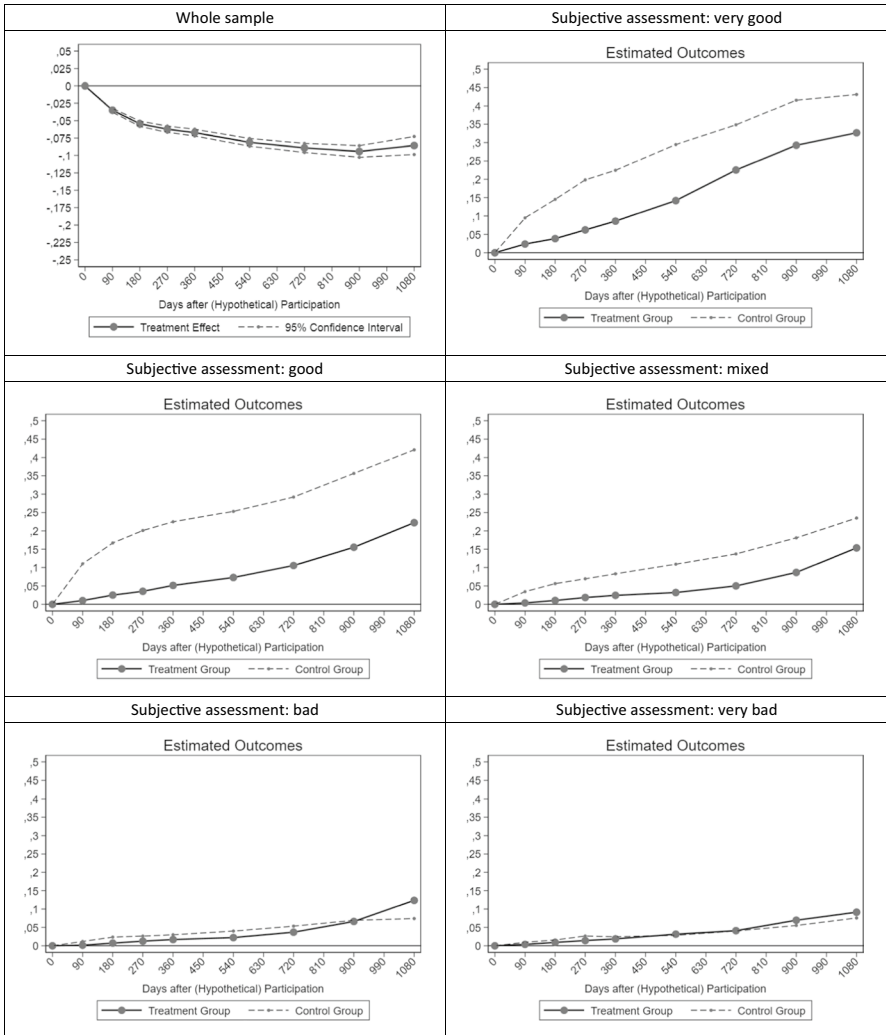


Fig. 9 Estimated potential outcome means for treatment and control group depending on assessment of the counsellor from the local employment agency. The upper left panel shows the ATT for the whole sample. The five remaining panels show the estimated potential outcome means, starting with the most favourable one (upper right corner) to the least optimistic one (lower right corner). Source: Integrated Employment Biographies (V11.01.00), own calculations

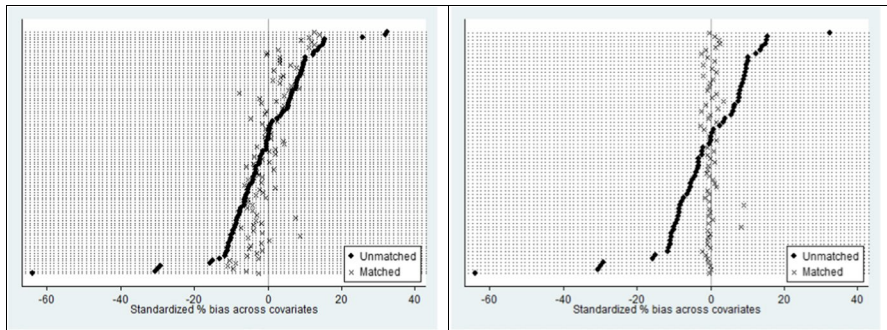


Fig. 10 Standardized bias before and after matching for radius matching (left panel) and entropy balancing (right panel) for the estimated effects 360 days after (hypothetical) programme start, register and survey data. Source: Integrated Employment Biographies (V11.01.00), own calculations

Funding Open Access funding enabled and organized by Projekt DEAL. Funding was provided by Bundesministerium für Arbeit und Soziales (DE).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahmad N, Svarer M, Naveed A (2019) The effect of active labour market programmes and benefit sanctions on reducing unemployment duration. *J Lab Res* 40:202–229
- Arendt JN, Kolodziejczyk C (2019) The effects of an employment bonus for long-term social assistance recipients. *J Lab Res* 40:412–427
- Arni P, Lalive R, van Ours JC (2013) How effective are unemployment benefit sanctions? Looking beyond unemployment exit. *J Appl Economet* 28:1153–1178
- Bellmann L, Caliendo M, Tübbicke S (2018) The post-reform effectiveness of the new German start-up subsidy for the unemployed. *Labour* 32(3):293–319
- Biewen M, Fitzenberger B, Osikominu A, Paul M (2014) The effectiveness of public-sponsored training revisited: the importance of data and methodological choices. *J Law Econ* 32(4):837–897
- Brown AJG, Koettl J (2015) Active labor market programs—employment gain or fiscal drain? *IZA J Labor Econ* 12(4).
- Caliendo M, Hogenacker J (2012) The German labor market after the great recession: successful reforms and future challenges. *IZA J Eur Labor Stud* 3(1).
- Caliendo M, Hujer R, Thomsen SL (2004) Evaluation der Eingliederungseffekte von Arbeitsbeschaffungsmaßnahmen in reguläre Beschäftigung für Teilnehmer in Deutschland. *J Labour Market Res* 37(3):211–237

- Caliendo M, Hujer R, Thomsen SL (2005) Sectoral heterogeneity in the employment effects of job creation schemes in Germany. *Jahrbücher Für Nationalökonomie Und Statistik* 226:139–179
- Caliendo M, Hujer R, Thomsen SL (2008) Identifying effect heterogeneity to improve the efficiency of job creation schemes in Germany. *Appl Econ* 40(9):1101–1122
- Caliendo M, Mahlstedt R, Mitnik AM (2017) Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies. *Labour Econom* 46:14–25
- Card D, Kluge J, Weber A (2010) Active labour market policy evaluations. A meta-analysis. *Econ J* 120:452–477
- Card D, Kluge J, Weber A (2018) What works? A meta analysis of recent active labour market programme evaluations. *J Eur Econ Assoc* 16(3):894–931
- Carling K, Richardson K (2004) The relative efficiency of labour market programs: Swedish experience from the 1990s. *Labour Econ* 11:335–354
- Chen S, Wang L (2020) SNAP participation, diet quality, and obesity: robust evidence with estimation techniques without external instrumental variables. *Empir Econ*. <https://doi.org/10.1007/s00181-020-01902-7>
- DeLoach SB, Kurt M (2013) Discouraging workers: estimating the impacts of macroeconomic shocks on the search intensity of the unemployed. *J Lab Res* 34:433–454
- Dengler K (2015) Effectiveness of sequences of One-Euro-Jobs for welfare recipients in Germany. *Appl Econ* 47(57):1–21
- Dorner M, Heining J, Jacobebbinghaus P, Seth S (2010) The sample of integrated labour market biographies. *J Context Econ* 130(4):599–608
- Fervers L (2019) Healing or deepening the scars of unemployment? The impact of activation policies on unemployed workers. *Work Employ Soc*. <https://doi.org/10.1177/2F0950017019882904>
- Frölich M, Heshmati A, Lechner M (2004) A microeconomic evaluation of rehabilitation of long-term sickness in Sweden. *J Appl Economet* 19:375–396
- Gerfin M, Lechner M (2002) A microeconomic evaluation of the active labour market policy in Switzerland. *Econ J* 112(482):854–893
- Hainmueller J (2012) Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. *Polit Anal* 20:25–46
- Heckman JJ, Hotz VJ (1989) Choosing among alternative nonexperimental methods for estimating the impact of social programs: the case of manpower training. *J Am Stat Assoc* 84(408):862–874
- Heyer G, Koch S, Stephan G, Wolff J (2012) Evaluation der aktiven Arbeitsmarktpolitik: Ein Sachstandsbericht für die Instrumentenreform 2011. *J Labour Market Res* 45:41–62
- Hohmeyer K (2012) Effectiveness of one-euro-jobs: do programme characteristics matter? *Appl Econ* 44(34):4469–4484
- Hohmeyer K, Wolff J (2012) A fistful of euros: is the German one-euro job workfare scheme effective for participants? *Int J Soc Welf* 21:174–185
- Huber M, Lechner M, Wunsch C, Walter T (2011) Do German welfare-to-work programmes reduce welfare dependency and increase employment? *German Econ Rev* 12(2):182–204
- Huber M, Lechner M, Wunsch C (2013) The performance of estimators based on the propensity score. *J Econ* 175:1–21
- Hujer R, Thomsen SL (2010) How do the employment effects of job creation schemes differ with respect to the foregoing unemployment duration? *Labour Econ* 17:38–51
- Hujer R, Zeiss C (2005) Macroeconomic impacts of job creation schemes on the matching process in West Germany. *Appl Econ Q* 51(2):203
- Imbens GW, Rubin DB (2015) *Causal inference for statistics, social, and biomedical sciences: an introduction*. Cambridge University Press, Cambridge
- Institut für Angewandte Wirtschaftsforschung (IAW) and Institut für Sozialforschung und Gesellschaftspolitk (ISG) (2015) Evaluation der Modellprojekte „Bürgerarbeit“ Endbericht zum 30. Juni 2015, Tübingen/Köln/Berlin. http://www.iaw.edu/tl_files/dokumente/IAW_ISG_Endbericht%20Buergerarbeit_Langfassung_final.pdf
- Kluge J (2010) The effectiveness of European active Labor market programs. *Labour Econ* 17:904–918
- Kraus F, Puhani PA, Steiner V (2004) Do public works programs work? Some unpleasant results from the East German experience. Some unpleasant results from the East German experience. In: Polaschek S (ed) *Research in labour economics*. JAI Press.
- Lechner M (1999) Earnings and employment effects of continuous off-the-job training in East Germany after unification. *Journal of Business & Economic Statistics* 17(1):74–90
- Lechner M, Wunsch C (2009) Active labour market policy in East Germany. *Econ Transit* 17(4):661–702

- Lechner M, Wunsch C (2013) Sensitivity of matching-based program evaluations to the availability of control variables. *Labour Econ* 21:111–121
- McCarthy I, Millimet D, Tchernis R (2014) The `bmt` command: methods for the estimation of treatment effects when exclusion restrictions are unavailable. *Stand Genomic Sci* 14(3):670–683
- Millimet DM, Tchernis R (2013) Estimation of treatment effects without an exclusion restriction: with an application to the analysis of the school breakfast program. *J Appl Economet* 28:982–1017
- Möller J (2015) Did the German model survive the labor market reforms? *J Labour Market Res* 48:151–168
- Mosthaf A, Schank T, Schwarz S (2021) Do supplementary jobs for welfare recipients increase the chance of welfare exit? Evidence from Germany. IZA Discussion Paper No. 14268. <http://ftp.iza.org/dp14268.pdf>.
- OECD (2020) Main economic indicators database. <https://doi.org/10.1177%2F0950017019882904>. Accessed 30 June 2020
- Rinne U, Zimmermann K (2012) Another economic miracle? The German labor market and the great recession. *IZA J Labor Policy* 1(3):1–21
- Rubin DB (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 66(5):688–701
- Sianesi B (2008) Differential effects of active labour market programs for the unemployed. *Labour Econ* 15(3):370–399
- Spermann A (2015) How to fight long-term unemployment: lessons from Germany. *IZA J Labor Policy* 4(1):1–15
- Statista (2021) Anzahl der Langzeitarbeitslosen in Deutschland im Jahresdurchschnitt 2008 bis 2021. <https://de.statista.com/statistik/daten/studie/666199/umfrage/anzahl-der-langzeitarbeitslosen-in-deutschland/>
- Thomsen SL, Walter T (2010) Temporary extra jobs for immigrants: merging lane to employment or dead-end road in welfare? *Labour* 24:114–140
- Wolff J, Stephan G (2013) Subsidized work before and after the German Hartz reforms: design of major schemes, evaluation results and lessons learnt. *IZA J Labor Policy* 2(19).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.