

Fervers, Lukas; Brändle, Tobias; Webel, Carina

Conference Paper

Decomposing a Matching Estimation - Correlated Unobserved non-random Selection, Substitution, and Causal Effects in a Large Scale ALMP Program in Germany

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Active Labour Market Policies, No. D20-V2

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Fervers, Lukas; Brändle, Tobias; Webel, Carina (2014) : Decomposing a Matching Estimation - Correlated Unobserved non-random Selection, Substitution, and Causal Effects in a Large Scale ALMP Program in Germany, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Active Labour Market Policies, No. D20-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100500>

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.

Decomposing a Matching Estimation

Correlated Unobserved non-random Selection, Substitution, and Causal Effects in a Large Scale ALMP Program in Germany¹

Tobias Brändle²

Lukas Fervers³

Carina Webel⁴

Friday, 14 February 2014

Preliminary version. Do not cite without permission!

Abstract

We analyse the treatment effects of a large scale German active labour market program. *Modellprojekte Bürgerarbeit* is a modified public employment scheme, with an *activation period* as a first step. Using data from individual employment records we perform matching estimations to analyse the treatment effect on the treated. In the cross-section estimation, we find reemployment rates to about 10%-points higher in the treatment group. Additionally, we consider that this effect might be upward biased due to correlated time-invariant unobserved heterogeneity and negative spill-over effects on non-participants. Therefore, we combine matching with difference-in-difference estimation (to control for unobserved heterogeneity). Doing so, we estimate a slightly lower treatment effect (about 8%-points). Finally, we use untreated individuals from regions with treated individuals (pseudo-treatment), and compare them to untreated individuals from regions without treatment. The negative pseudo-treatment effect lies around 2%-points, indicating negative spill-over effects. We conclude that the 10%-point difference from the matching estimation can be decomposed into a causal effect (6%-points), correlated unobserved heterogeneity and negative spill-over effects (2%-points each).

Keywords: active labour market program, public employment scheme, propensity score matching, difference-in-differences, spill-over effects, integrated employment biographies

JEL Code: J68, J14

¹ This project has been realised with financial support from the German Ministry of Labour and Social Affairs, but does not necessarily reflect its opinions and views. The data has been provided by the Institute for Employment Research in Nürnberg, with a special thanks to Steffen Kaimer. We gratefully acknowledge the collaboration of Bernhard Boockmann, Katrin Harsch, Andrea Kirchmann, Hans Verbeek, Regina Weber, Helmut Apel, Dennis Egenolf, Michael Fertig, Martin Rosemann, and Marian Weimann. Thanks also to Johannes Binder for valuable research assistance. All errors remain ours.

² **Corresponding Author;** Institute for Applied Economic Research (IAW) Tuebingen; tobias.braendle(at)iaw.edu.

³ Institute for Applied Economic Research (IAW) Tuebingen; lukas.fervers.(at)iaw.edu.

⁴ Chamber of Employment Saarland; carina.webel(at)arbeitskammer.de

1. Introduction

In recent decades, active labour market programmes (ALMPs) have gained increasing importance in almost all OECD countries. Industrialised countries spend large amounts of funds to facilitate labour market re-integration of the unemployed. In economic policy analysis, ALMPs are regarded as an important means to improve labour market performance. At the same time, the question of whether ALMPs actually foster labour market reintegration is a debated issue in the economics literature. In this research project, we provide empirical evidence for one of the most recent ALMPs in Germany, the so called *Modellprojekte Bürgerarbeit*. It is a modified public employment scheme, consisting of two periods. The first one supplies the unemployed with intensified counselling services and access to training opportunities. Only if they participate in this *activation period*, and if they do not find a job during that time, they qualify for the second period, which is a public employment scheme. The Institute of Applied Economic Research (IAW) performs the evaluation of the programme from 2010 to 2015, the latest interim report can be found at IAW, ISG (2013).

In this paper we present results from an analysis of the effect of participation in the *activation period* on labour market re-integration. The innovation of our paper is twofold. On the one hand, the programme is of very large scale, so identifying its causal effect on participants is highly relevant from a policy-making point of view. The other aspect is a more methodological one, since the institutional design of the programme involves high challenges when it comes to the identification of the treatment effect: In spite of its large scale, job centres can voluntarily participate in the programme. Moreover, for participating job centres, the method of selecting participants for the programme, i.e. defining the target group, can be chosen freely. Additionally, in most participating job centres, the unemployed belonging to the target group were allowed to decide voluntarily about participating in the programme. As a consequence, it is possible that participating and non-participating unemployed individuals coexist in a participating job centre. Therefore, identifying the causal effects of the programme is non-trivial since non-random selection into treatment might happen at the level of the job centre as well as at the individual level. Moreover, spill-over effects are likely to occur (see Hujer et al 2009). At the same time, the institutional design opens up opportunities to account for non-random selection as well as possible spill-over effects on untreated individuals. Given the availability of high quality panel, we can decompose descriptive comparisons and cross-section matching estimation, which might suffer from a selection bias and cannot account for spill-over effects, into the actual causal effect, correlated unobserved heterogeneity and possible spill-over effects on non-participants. To do so, we start with a simple matching-estimation to account for observable differences between participants and non-participants (see, e.g. Caliendo and Hujer, 2006). Additionally, we combine the matching estimation with a difference-in-differences design to account for time-invariant unobserved heterogeneity. Finally, we consider possible substitution effects, i.e. a spill-over effect of the programme on non-participants. To estimate this spill-over effect, we compare untreated individuals from participating job centres (which might suffer from a spill-over effect, i.e. the pseudo-treatment group) with individuals from non-participating job centres (which are by definition unaffected).

The remainder of this paper is structured as follows. A short review of the literature on ALMP evaluation, with a focus on methodological difficulties builds Chapter 2. A summary of the programme *Modellprojekte Bürgerarbeit* will follow, accompanied by a presentation of the data and descriptive statistics in Chapter 3. We will recap the empirical methods used in this paper in Chapter 4. The results are presented in Chapter 5 alongside discussion of further steps to be taken. Chapter 6 concludes.

2. Literature Review

Due to the extension of active labour market policies (ALMPs) in western OECD countries, there has been an increasing interest in the effects of such programmes. As a result, there is a huge and rapidly growing body of microeconomic literature which analyses the effects on the participants in terms of exit from unemployment and post-unemployment employment outcomes (i.e. employment stability, wages and type of contract). Given the huge variety of programme types, it has become common standard to distinguish between training programmes, counselling services and public employment schemes. Even though the results differ somewhat with regard to country sample and time span under observation, there are three main conclusions that can be drawn. First, training programmes reduce exit from unemployment in the short run, but improve employment outcomes in the long run. This pattern has been confirmed for Switzerland (Arni 2010), East and West Germany (Lechner et al. 2007; Lechner et al. 2011), Denmark (Munch and Skipper 2008), France (Crépon et al. 2012), Sweden (Larsson 2003), Norway (Raaum et al. 2002) and Romania (Rodriguez-Blanas 2010). Additionally, Cavacao et al. (2004) identify a positive long-run effect of training programmes on the probability of getting a regular contract. Stephan and Pahnke (2011) further distinguish between different durations of training programmes. They report that short programmes are more effective in the short run, while this effect is exceeded in the long run by longer training programmes. The results concerning counselling services are less dependent on the duration of the time span under observation. For different time spans, positive results have been found for the Netherlands (Kastoryano and van der Klaauw 2011), Switzerland (Arni 2010), Romania (Rodriguez-Blanas 2010) and France (Behaghel et al. 2011). The effects of public employment schemes are less consistent. While some studies find positive effects (Caliendo et al 2008), they overall tend to worsen employment outcomes, which is probably caused by locking-in and/or stigma effects (Card et al 2010). Due to the increasing number of evaluations, two meta-analyses have been conducted by Card et al. (2010) and Kluve (2010). These two meta-analyses confirm the positive effect of counselling services and training programs (in the long run), as well as the negative effect of public employment schemes. Finally, positive effects of tighter job search requirements on exit from unemployment have been detected in the Netherlands (Bloemen et al. 2011) and Dyke et al. (2006).

Methodologically, these non-experimental evaluations centre on the problem of non-random selection into these programmes, which cause a selection bias when the treatment and control group are compared. This problem is solved by matching or regression (to account for non-random selection with regard to observable covariates), and/or Difference-in-Differences, Instrumental Variable estimation or a regression discontinuity design (when the non-random selection occurs with regard to unobservable variables). At the same time, controlling for

possible spill-over-effects on non-participants is often still an unsolved problem. Existing evaluations are thus often based on the controversial stable unit treatment value assumption (SUTVA). Since some ALMPs are likely to display substitution effects, existing evaluations which find positive effects might suffer from an upward bias. Methodologically, the main innovation of this research project is that we can exploit the institutional setting to estimate possible spill-over effects.

3. Programme Overview, Data, and Descriptive Statistics

3.1 The Structure of the Program and Potential Sources of Bias

General Structure

The active labour market programme *Modellprojekte Bürgerarbeit* is run by the Ministry of Labour and Social Affairs and has started on July 1st, 2010 in 197 job centres across Germany, which translates into coverage rate of about 50 %. The funds for the project are evenly split by the federal budget and the European Social Fund, each supplying just over 200 million Euros a year until the end of 2014. The programme consists of two periods, the activation period and the employment period. The activation period, which this paper focuses on, lasts at least six month for each participant.

Job centres were allowed to define the characteristics of potential participants, i.e. the target group, themselves. The only guideline given by the Ministry of Labour and Social affairs concerning the selection of participants into the programme was that they had to be registered unemployed as stated in § 16 SGB III. Job centres were allowed to decide freely to select individuals out of the whole group of unemployed or to define certain characteristics to focus on (e.g. single parents, a certain range of age, certain results of profiling, migrants). Participating job centres differ widely in the definition of their target group. The selection of participants out of the given target group was performed by the job advisers. In most job centres, selected unemployed were asked by their advisers if they would like to take part in the programme or not and could voluntarily decide about participating in the programme.

Activities during the activation period differed widely between job centres, not many guidelines were given by the Ministry. In the majority of the job centres, a central element of the activation period was a more intense and regular contact between job advisers and participating unemployed. Apart from performing the activation period themselves, job centres also had the possibility to commission external providers of social services to do the activation period.⁵

A requirement for taking part in the employment period of the programme was a completion of the activation period. Only then, and if they did not find a job until then, the participants could get assigned to the employment phase, which can last up to 36 months. The aim of the programme's second part is to provide for about 34,000 public employment positions with a focus on integrating participants into the first labour market during or after the programme. It

⁵ The IAW conducted a survey in all job centres as well as several case studies to find out about the details of the different designs of the activation period in job centres.

is important to know that additional funding was only provided for this second period of the programme and not for the activation period. If job centres were not able to acquire additional funding from other sources such as the European Social Fund on the state-level or the local budget, they had to organise the activation period under the constraints of their regular budget. This was the case for almost 60% of all participating job centres.

Sources of Potential Bias

In order to get an idea about possible selection bias and spill-over effects, it is crucial to understand the assignment mechanism. Due to the institutional context, there are two sources of selection. First of all, job centres participate voluntarily. As a consequence, it is possible that on the one hand, job centres confronted with a more severe labour market situation have a higher probability of participating because they are in more need of new approaches to integrate the unemployed into the first labour market. On the other hand, it might be conceivable that job centres with a better regional labour market situation have a higher propensity for participating because they have more resources to provide for the activation. Additionally, within the job centres there are participants and non-participants.

At the level of the job centres, we use information from an email survey and official regional statistics to get insights on selection into participation. There is no evidence for a positive selection in the data. If anything, there seems to be a slightly negative selection. Participation rates are higher in East Germany (which is economically weaker), and average unemployment rates before the start of the programme have been higher in the regions of the participating job centres. Moreover, it is worth noticing that the change of the regional unemployment rates was exactly the same in the regions with participating and not participating job centres in the year before the program started (2009 to 2010, -7.5%, IAW 2012). Apparently, there is no evidence for an *Ashenfelter's dip* at the level of the job centres. These data are in line with the results from the survey, which indicate that the main reason for job centres to participate was the possibility to create new job opportunities for hard-to-place workers and that this was especially the case in job centres with a problematic regional labour market situation.

A comparison of participants and non-participants at the individual level reveals a different picture. We use information on the assignment of unemployed individuals into the programme by the job centres. It shows that assignment is very heterogeneous: In some job centres participation is voluntary, in others it is compulsory and non-participation may be sanctioned.⁶ A number of case studies accompanying the quantitative evaluations show very different situations to be found regarding the perceived success of programme participation.⁷ In summary, according to the job centre employees, the potential success of the programme largely depends on the choice of participants and on their motivation to find a job in the first labour market (and not just to entering the public employment scheme).

⁶ There are reports from a number of cases of non-compliance. Participants call in sick, shirk in meeting the programme targets, or even drop out of welfare support. In some cases participants immediately find a job after being placed into the programme. These are suspected of having worked undeclared before.

⁷ While we refer to "programme participation" in general, most of the results of the evaluation so far focus on the activation period.

We have no detailed knowledge about the exact assignment mechanism in each job centre, but the results from our first stage of the propensity score matching indicate that unemployed workers with favourable characteristics had a higher participation probability. For example, it is higher for males, prime-age workers and non-migrants. At the same time, having health problems, chronic illnesses or children below the age of three lowers the chance to get assigned to the program. This indicates a positive selection bias into the programme at the level of the individuals. However, in most job centres the officials cite a better rate of contacts with the participants as a main driver of higher integration rates. Therefore, the majority of job advisers have high expectations from the *activation period*, which should provide exactly this. As a consequence, after the end of the *activation period*, most job centres use internal evaluation, because they would like to improve their work and to assess if the programme has been a success for the participants and how to place them.

Finally, it is worth thinking about possible spill-over effects. The general problem of substitution is as present as in any other ALMP programme. It is possible that the intensified counselling services do not lead to net job growth at the aggregate level, but the participants just find them more quickly and therefore substitute the non-participants that might have found the job otherwise. Moreover, the financing mechanism raises concerns about possible spill-over-effects. In most cases, the job centres did not receive an additional budget to finance the activation period, but they had to rely on the regular resources. Accordingly, the vast majority (87%) of the participating job centres did not hire additional staff for the programme. Under given budget constraints, this is likely to result in a redistribution of resources in favour of the participants, i.e. a reduction in spending and effort for the non-participants. Positive employment effects on participants might thus be upward-biased. All in all, this institutional setting raises concerns about positive selection at the individual level as well as possible spill-over effects.

3.2 Data and Descriptive Statistics

Integrated Employment Biographies

The evaluation is based on specific samples from the Integrated Employment Biographies (IEB), an administrative data set assembled and provided by the German Federal Employment Agency for the evaluation of most German ALMPs. These data consists of several different datasets collecting information on employed and unemployed individuals in Germany. We have information on employment spells and rudimentary individual characteristics from *Beschäftigtenhistorik* dataset. These can be matched with information on more detailed information on unemployed workers looking for a job in the *Arbeitsuchendenstatistik*, including information on the last job performed and qualifications. Then, there is additional insight on the household and further characteristics relevant for the receipt of unemployment benefits in the *Leistungsempfängerhistorik*. Finally, we know for each individual the past and current active labour market programmes, trainings or further assistance received by the job centres in the *Maßnahmeteilnahmehistorik*. We can observe a large number of control variables and are able to track all observations for several years. Detailed information on a random sample from the IEB can be found in Dorner et al. (2010). A discussion of the data

and their use in the evaluation of ALMPs is made by Lechner and Wunsch (2013), who we follow broadly in the preparation of the data.

For the evaluation, we use information on participating individuals entering unemployment between 2010 and 2011 and compare them to non-participants both contemporarily and to before the programme has been established. The non-participants are drawn both from job centres that participate in the programme and ones which do not. We track the individuals' employment outcomes over a period of 12 to 21 months after entering the programme (for participants) or after a cut-off date, which is the start of the programme on July 1st, 2010 (for non-participants). For the difference-in-differences sample we also track individuals from the respective sample following July 1st, 2009, one year before the start of the programme. The dataset consists of five samples: one for 69,000 participants further described below and four for non-participants each covering a certain control group and covering 125,000 individuals. To enter the sample, individuals must require two facts: they have to be dependent on means-tested public welfare benefits for individuals able to work (UB II), i.e. they are long-term unemployed, and they have to be unemployed and looking for work at the time they enter their welfare spells.

As outcome variables we use integration into the first labour market and exit from public welfare dependence. The first labour market is defined as having a job subject to social security that does not involve being in an assistance programme of any sort.⁸ We do not distinguish between part-time and full-time, as there is only imprecise information on that in the data, and we do not directly condition on earning a certain wage.⁹ Public welfare dependence is measured as receiving means-tested benefits (UB II). While being employed can be one reason to exit public welfare, there are also other. The integration of a household member, esp. of spouses can lead to a suspension of welfare benefits. Also, a change in the household composition, e.g. leaving children or marrying can do so. In addition, individuals may be declared unable to work, or may leave the welfare system to being out of the labour force for their own reasons. Therefore, we think both variables complement each other very well.

Programme Participation

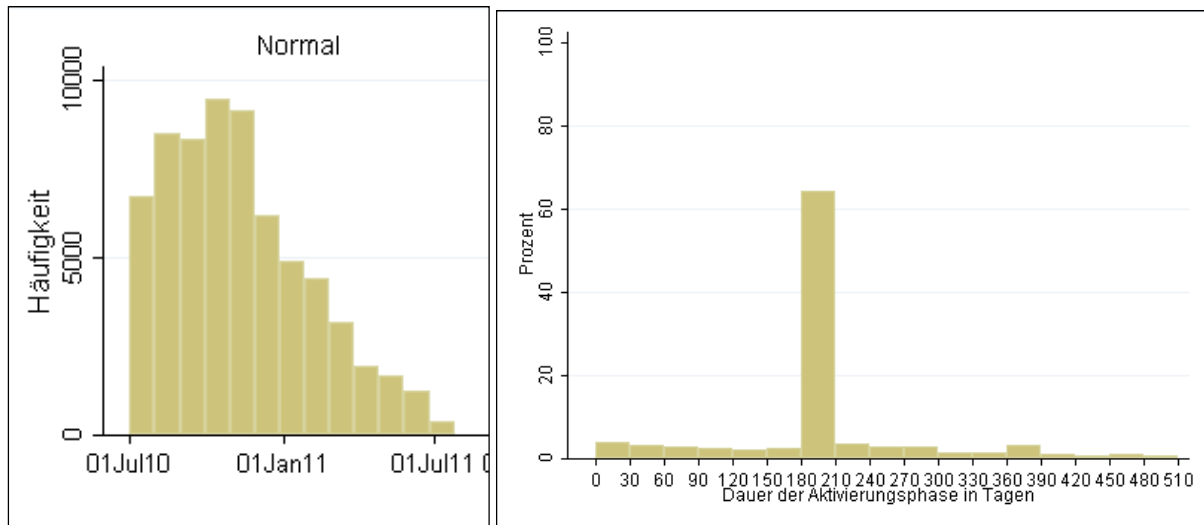
Programme participation for the *activation period* is usually not part of the IEB. Therefore, the job centres have exceptionally collected information on participation based on internal documentations between July 1st, 2010 and June 30th, 2011. Using this approach about 69,000 individuals can be observed to have participated in the *activation period*.¹⁰

⁸ Formally, the employment spell must not collide with an entry in the Maßnahmeteilnahmehistorik, meaning that it is not subsidized publicly nor part of a public employment scheme or any other form of active labour market programme.

⁹ However, the wage must implicitly be as high as to be earning more than 400 Euros a month, i.e. above marginal employment.

¹⁰ Among others this procedure leads to some potential inconsistencies, especially regarding individuals entering the program after June 30th, 2011. Furthermore, for a number of cases (about 17,000) the collection of the data is terminated before the actual end of the activation phase, such that this information is to be treated with caution.

Figure 3.1: Begin and Duration of the Activation Period



Source: Own calculations based on samples of the Integrated Employment Biographies.

Programme participation involves participation in the activation period on the first stage. The job centres have to plan how many unemployed to activate in order to be able to fill the number of public employment jobs afterwards (plus some overhead). In a number of the cases this has proven to be difficult as the possibility of unemployed entering the first labour market during or after the *activation period* leads to the need of “post-activation”, i.e. increasing the number of participants over time. For these people, it can be the case that they have not been in the programme for the minimum amount of six months before entering the public employment scheme.

Figure 3.1 displays these circumstances by displaying the beginning of the *activation period* in the left panel and its duration in the right panel. You can see that a large number of unemployed start participation well over half a year after the official start of the program, which is July 15th, 2010. Furthermore, only about 65 % of participants stay in the *activation period* for the planned duration of six months, while both about 15 % of participants leave early (among those not only individuals that have been integrated into the labour market) or stay longer.¹¹

4. Methods

From a methodological point of view, the descriptive difference between treatment and control group can result from three sources. First of all, it can result from a selection bias, i.e. differences between treatment and control group which are correlated with program participation probability and employment outcomes. Secondly, the program can have a direct causal effect on employment outcomes, as originally intended by policy-makers. Finally, the program might exert negative spill-over effects on non-participants in the same local labour market (violation of the stable unit treatment value assumption (SUTVA)). While the general notion of spill-over effects is not completely new, it is under researched and of particular importance in this setting: Since job centres finance the program with their regular budget,

¹¹ It is also to note that a large number of individuals seem to have artificial end dates, for example the end of the calendar year or dates that are situated several years into the future.

they are (under given budget constraints) likely to reduce spending and effort for non-participants. Identifying the treatment effect (which is the actual parameter of interest) requires decomposing the observed difference into these three parts.

To do so, we proceed in three steps. First, we perform a matching estimation based on propensity score matching. Performing a matching estimation requires choosing between different matching algorithms. This choice is non-trivial, since it involves a trade-off between bias and efficiency. While evidence from Monte-Carlo-Simulation on the performance is inconclusive, we follow the theoretical idea that our exceptionally large dataset suggests a stronger focus on bias reduction than efficiency. We thus choose nearest-neighbour matching with comparatively few control observations ($k=4$). Additionally, we introduce a rather tight calliper (difference $> 0.01\%$), which excludes observations for which the next neighbour is too far away. Standard balancing tests show that the performance of propensity score matching is very satisfactory. Both mean and median bias between treatment and control observations for the covariates are reduced substantially by matching. Re-estimation of the Probit model on the matched sample results in an explained treatment variation of almost zero, as measured by the McFadden- R^2 and the corresponding LR statistic (Sianesi, 2004). This estimator successfully takes observable non-random selection into account, but relies on the conditional independence assumption (CIA, i.e. absence of correlated unobserved heterogeneity) and the stable unit treatment value assumption (SUTVA, i.e. absence of spill-over effects). Therefore, we go on by combining matching approaches with difference-in-differences (DiD) estimation in order to refute concerns about time-invariant correlated unobserved heterogeneity. We regard this step as crucial, since treatment and control group might differ in terms of unobservable factors such as motivation. Finally, we need to find a way to estimate the magnitude of the spill-over effects. In contrast to the solution of self-selection problems, standard procedures are still biased in case of violations of the SUTVA (cf. Imbens and Wooldridge, 2009). Here, we exploit the institutional setting to estimate the hypothesized spill-over effect: we perform a DiD-estimation with pre-matching in which we use non-participants in participating job centres (who might suffer from negative spill-over effects) as a treatment group, and non-participants in non-participating job centres (who can be assumed not to be affected in any way by the program) as control group. The difference between these groups can then be regarded as an approximation of the spill-over effect (cf. Hudgens and Halloran, 2008).

5. Results and Discussion

In this section we present the results of our quantitative analysis. First, we present matching estimates from participants compared to non-participants in the same job centre. Then, we proceed by showing results from difference-in-differences estimation regarding a) the effects of programme participation compared to control groups from non-participating job centres and b) the spill-over effects of the programme estimated by comparing non-participants on participating and non-participating job centres. The results are followed by a discussion.

5.1 Matching Estimation

We start by presenting results from the propensity score matching. First stage estimates and balancing tests are discussed in Chapter 4 and available upon request. Therefore, we will

focus on displaying the average treatment effect on the treated (ATT). It measures the difference in outcome probabilities (to enter employment in the first labour market or to exit public welfare) between programme and control group conditioned on equal propensity scores.

The results are displayed in Figure 5.1. The average ATTs are measured in percentage point differences and in days after programme participation (with the confidence interval around it for different points in time).

Figure 5.1: Treatment Effects of Participation in the Activation Period, Matching Estimation



Source: Own calculations based on samples of the Integrated Employment Biographies.

The left panel of Figure 5.1 shows the ATT on integration into the first labour market. It can be seen that it is close to zero during the duration of the *activation period* itself, which is to last for 180 days. However, after that the ATT quickly rises to become positive and significant at an eight percentage point margin. It remains at this level for most of the time afterwards.¹² Regarding exit from public benefits, the right panel of Figure 5.1 shows another picture. There is no positive treatment effect from participating in the *activation period* at any point in time. Indeed, there are small but significant negative ATTs during the *activation period* itself and starting again one year after programme participation. The size of the effect reaches up to 3%-points.

ATTs of 8%- and 3%-points might seem small. However, regarding the overall integration or exit probabilities of the individuals in our sample, you might reconsider this. Figure A.1 in the Appendix shows on average and one year into the sample integration and exit rates of about 25%. Therefore, the absolute numbers indicate that the ATTs are of substantial size, such that the 8%-points difference in integration rates corresponds to a relative difference of about 40% and the 3%-points difference in exit rates corresponds to a relative difference of about 15%.

We conclude that directly after the end of the initial *activation period* the integration probability of treated individuals is much higher. Therefore, programme participants profit very strongly regarding their chances of getting a job in the first labour market. However,

¹² There is another slight increase at the end of the observation period, when the ATT reaches 10%-points. However, due to a decrease in the number of observations, this should be interpreted with caution.

treated individuals do not exit public welfare at a higher rate. This is especially true during the *activation period* itself, but also one year after the start of the treatment. The causes of these differentiated effects are subject for further work.

Nevertheless, these differences are substantially smaller compared to naive descriptive comparisons, indicating that non-random selection into program participation plays a crucial role. Therefore, we proceed to alternative estimation techniques in the following.

5.1 Conditional Difference-in-Differences

Matching itself does not overcome all possible sources of bias in estimating treatment effects. First, the matching results are based on a cross-section analysis, such that unobserved heterogeneity might play an important role. Second, the causal effects of the *activation period* could not only result from a better integration of treated individuals, but from a worse one for the control group, e.g. through reducing spending and effort for non-participants. We tackle both problems using conditional difference-in-differences methods.¹³

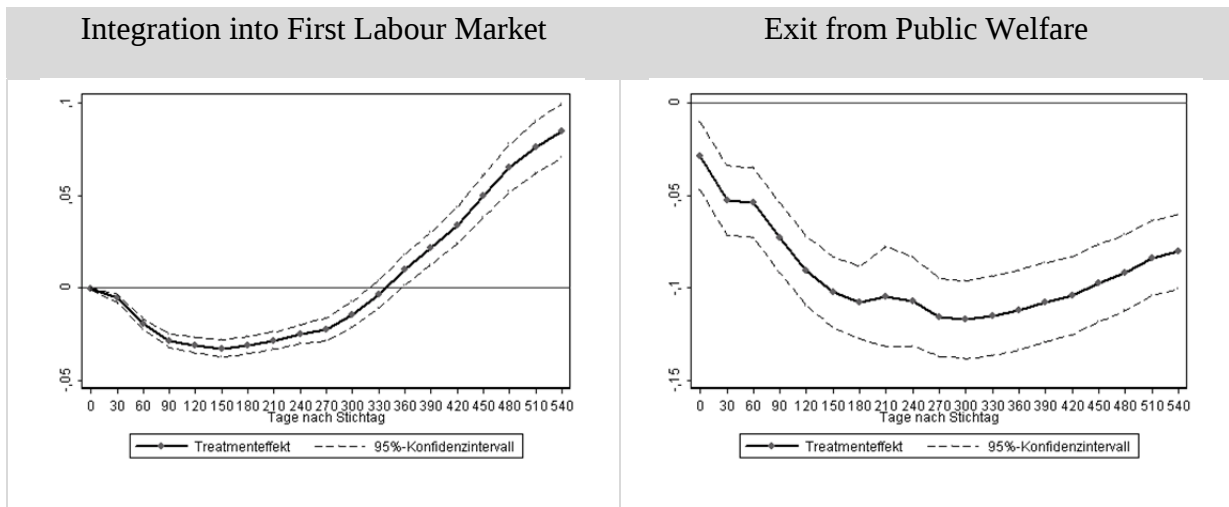
First, we estimate a treatment effect of programme participation based on treated individuals by comparing them to untreated individuals in participating job centres before the start of the programme (DiD-estimation 1) and to untreated individuals in non-participating job centres before and after the (hypothetical) start of the programme (DiD-estimation 2).¹⁴ Using this approach we control for time-invariant unobserved heterogeneity. Second, we estimate a (non-)treatment effect for the non-participants in participating job centres, i.e. negative spillover effects. We therefore compare these untreated individuals with untreated individuals from non-participating job centres.

Figure 5.2 displays the results from the first approach. The average ATTs are measured in percentage point differences and in days after the start of the programme (compared to one year earlier). A 95%-Confidence interval is supplied as well.

¹³ DiD estimation rests on the common trends assumption. We have checked this. Results are available upon request. Especially for the first approach common trends seem to be violated. Therefore, the results have to be interpreted with caution.

¹⁴ Please note that our approach does not involve observing the same individuals over multiple time periods (panel analysis). This would make no sense since you would compare the same people with different unemployment durations. However, time since the last job is one of the most important determinants of labour market integration.

Figure 5.2: Treatment Effects of Participation in the Activation Period, Difference-from-Differences Estimation



Source: Own calculations based on samples of the Integrated Employment Biographies.

The left panel of Figure 5.2 presents results for integration into the first labour market, the right panel for exit from public welfare. Here, the interpretation is slightly different from the matching results. Regarding integration, there seems to be a negative effect during and after the *activation period*. This could be caused by lock-in-effects, but the results do not allow for a precise interpretation in that direction. Furthermore, the *activation period* usually does not consist of measures that have been shown to cause these kinds of effects in the past. More probably, the differences come from different measurement of timing. Here, the programme and the control group are measured at the same calendar date (unlike in the matching estimation, where the participants are measured from the date of their individual participation). Further work will explore these differences. However, after one year, the effects turn. At the end of the observation period they reach a positive ATT of about 8%-points, similarly to the matching estimation. Regarding exit from public welfare, the results seem qualitatively similar to the matching results: there is a negative treatment effect of programme participation. However, quantitatively the effect is much larger. It amounts for up to 12%-points.

Finally, we analyse the existence of spill-over effects, by using non-participants in participating job-centres as (pseudo-) treatment group, and non-participants in not participating job centres as control group. Figure 5.3 displays the results in a similar manner as the figures above.

Figure 5.3: Substitution Effect of Potential Activation on Labour Market Integration, Difference-from-Differences Estimation



Source: Own calculations based on samples of the Integrated Employment Biographies.

It can be seen that for integration into the first labour market the point estimates indicate a negative spill-over effect. However, the effect is very small and insignificant throughout. For exit from public welfare, we see negative pseudo-treatment effect. It is significant and reaches a coefficient size of up to 3%-points. This amounts to about a third of the negative ATT in the DiD and about all of the negative ATT in the matching estimation. Therefore, we conclude that spill-over effects might indeed play role in explaining the effects of the programme.

However, increasing the efficiency of the spill-over estimation is still a crucial methodological challenge, since it would result in a more precise identification.

5.2 Discussion

All in all, four main conclusions can be drawn from the analysis. First of all, the *activation period* tends to display weak locking-in effects. In none of the estimations, there is a positive effect on exit from unemployment during the programme. In the DiD estimation, it is slightly negative. Secondly, the program seems to have a substantially positive effect on employment outcomes in the long and medium run. Even if selection on observables and unobservables is taken into account, the average treatment effect approaches 8%-points. Thirdly, the difference between the descriptive comparison, the cross-section matching approach and the conditional DiD approach indicates that selection occurs on observables as well as on time-invariant unobservable variables. The remaining question mark is the change of the time dependence in the DiD-estimation. Finally, there seems to be weakly negative but insignificant spill-over effects on non-participants. If the points estimate are taken for granted, the 8%-points effect can be decomposed into a direct effect of about 8%, and a negative spill-over effect of 2% (at the end of the observation period).

6. Conclusion

In this paper, we have analysed the effect of the *activation period* of the ALMP project *Modellprojekte Bürgerarbeit* in Germany. We have performed different estimation approaches to identify the causal effect of the programme, by taking non-random selection

and negative spill-over effects into account. Our results can be summarized as follows. During the *activation period* itself, there is no significant difference in re-employment probabilities between treatment and control group. The absence of a positive effect is consistent with theoretical expectations, since the first period is likely to have locking-in effects, e.g. due to training participation. However, after six months in the *activation period* and up until one and a half year a simple matching estimation reveals the incidence of unemployed workers re-entering the labour market to be 10%-points higher in the treatment group. If matching is combined with DiD estimation, this reduces to 8%-points, which implies that the simple matching estimation might suffer from unobserved time-invariant selection bias. Nevertheless, it does not seem to be very strong compared to the overall effect. Finally, comparing non-participants in participating and non-participating job centres points to slightly negative but insignificant spill-over effects: The incidence of reemployed workers is about 2%-points smaller in the treatment group, albeit not significantly different from zero.

All in all, the first period of the program can be claimed to be successful. The employment effects are substantial. At the same time, it has to be considered that the costs of the program are likely not to be too high, since it is rather short. The detected locking-in effects are also weak, which is why a positive net effect in terms of cumulated time in employment will be achieved after a couple of months. We have not yet performed a cost-benefit analysis, but considering what has been just said, positive fiscal efficiency is likely to be achieved after a rather short period of time. Whether or not the second period of the program will also foster integration into non-subsidized employment will be subject to another evaluation project. After all, this evaluation indicates that long-term unemployed are not a lost group, but their successful re-integration depends on good or bad policies.

7. References

- Arni, P. (2010): How to Improve Labor Market Programs for Older Job-Seekers? Evidence from a Field Experiment, Ms., University of Lausanne.
- Behaghel, L., Crepon, B., and Gurgand, M. (2014): Private and Public Provision of Counseling to Job-Seekers: Evidence from a Large Controlled Experiment, *American Economic Journal: Applied Economics*.
- Bloemen H., Lammers M., and Hochgürtel S. (2011): Job Search Requirements for Older Unemployed: Transitions to Employment, Early Retirement and Disability Benefits. *DP 01/2011-002* (revised version October 2012).
- Dorner, M., Heining, J., Jacobebbinghaus, P., and Seth, S. (2010): Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) 1975-2008, *FDZ Datenreport*, 01/2010.
- Caliendo, M., Hujer, R., and Thomsen, S.L. (2008): The Employment Effects of Job Creation Schemes in Germany: a Microeconomic Evaluation, in: Millimet, D., Smith, J., and Vytlacil, E. (Eds): *Advances in Econometrics Vol. 2, Estimating and Evaluating Treatment Effects in Econometrics*, Elsevier: Amsterdam, 383–430.
- Caliendo, M., and Hujer, R. (2006): The Microeconomic Estimation of Treatment Effects - An Overview, *Journal of the German Statistical Society*, 90(1), 197-212.
- Card, D., Kluve, J., Weber, A. (2010): Active Labour Market Policy Evaluations. A Meta-Analysis, *The Economic Journal*, 120, 452–477.
- Cavaco, S., Fougère, D., and Pouget, J. (2004): Estimating the Effect of a Retraining Programme for Displaced Workers on their Transition to Permanent Jobs, *IZA Discussion Paper No. 1513*.
- Crépon, B., Ferracci, M., and Fougère, D. (2012): Training the Unemployed in France: How Does it Affect Unemployment Duration and Recurrence? *Annals of Economics and Statistics* 107-108, 175-199.
- Dyke, A., Heinrich, C.J., Mueser, P.R., Troske, K.R., and Jeon, K.S. (2006): The Effects of Welfare-to-Work Programme Activities on Labor Market Outcomes, *Journal of Labor Economics*, 24(3), 567–606.
- Hudgens, M., and Halloran, E. (2008): Toward Causal Inference with Interference, *Journal of the American Statistical Association*, 103(482), 832-842.
- Hujer, R., Rodrigues, P., and Wolf, K. (2009): Estimating the Macroeconomic Effects of Active Labour Market Policies using Spatial Econometric Methods, *International Journal of Manpower*, 30(7), 648-671.
- Institut für Angewandte Wirtschaftsforschung (IAW) and Institut für Sozialforschung und Gesellschaftspolitik (ISG) (2013): Evaluation der Modellprojekte „Bürgerarbeit“ Zwischenbericht zum 30. April 2013, publicly available under http://www.iaw.edu/w/IAWPDF.php?id=1017&name=IAW_ISG_2_Zwischenbericht_Buergerarbeit.pdf.

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

Kastoryano, S., and van der Klaauw, B. (2010): Dynamic Evaluation of Job Search Assistance, University of Amsterdam, VU University Amsterdam, Tinbergen Institute.

Kluve, J. (2010): The Effectiveness of European Active Labor Market Programs, *Labour Economics*, 17, 904–918.

Larsson, L. (2003): Evaluation of Swedish Youth Labour Market Programmes, *Journal of Human Resources*, 38(4), 891–927.

Lechner M., Miquel, R., and Wunsch, C. (2007): The Curse and Blessing of Training the Unemployed in a Changing Economy: the Case of East Germany after Unification, *German Economic Review*, 4(8), 468–507.

Lechner, M., and Wunsch, C. (2013): Sensitivity of Matching-Based Program Evaluations to the Availability of Control Variables, *Labour Economics*, 21, 111–121.

Lechner, M., Miquel, R., and Wunsch, C. (2011): Long-run Effects of Public Sector Sponsored Training in West Germany, *Journal of the European Economic Association*, 9(4), 742–784.

Munch, J.R., Skipper, L. (2008): Programme Participation, Labor Force Dynamics, and Accepted Wage Rates, in: Fomby, T., Hill, T.C., Millimet, D.L., Smith, J.A., Vytlačil, E.J. (Eds): *Modelling and Evaluating Treatment Effects in Econometrics (Advances in Econometrics, Volume 21)*, Emerald Group Publishing Limited, 197–262.

Raaum, O., Torp, H., and Zhang, T. (2002): Do Individual Program Effects Exceed the Costs? Norwegian Evidence on Long Run Effects of Labour Market Training, Memorandum 15: Department of Economics, University of Oslo.

Rodriguez-Planas, N., Benus, J. (2010): Evaluating Active Labour Market Programs in Romania, *Empirical Economics*, 38(1), 65–84.

Stephan, G., and Pahnke, A. (2011): The Relative Effectiveness of Selected Active Labour Market Programs: An Empirical Investigation for Germany, *The Manchester School*, 79(6), 1262–1293.

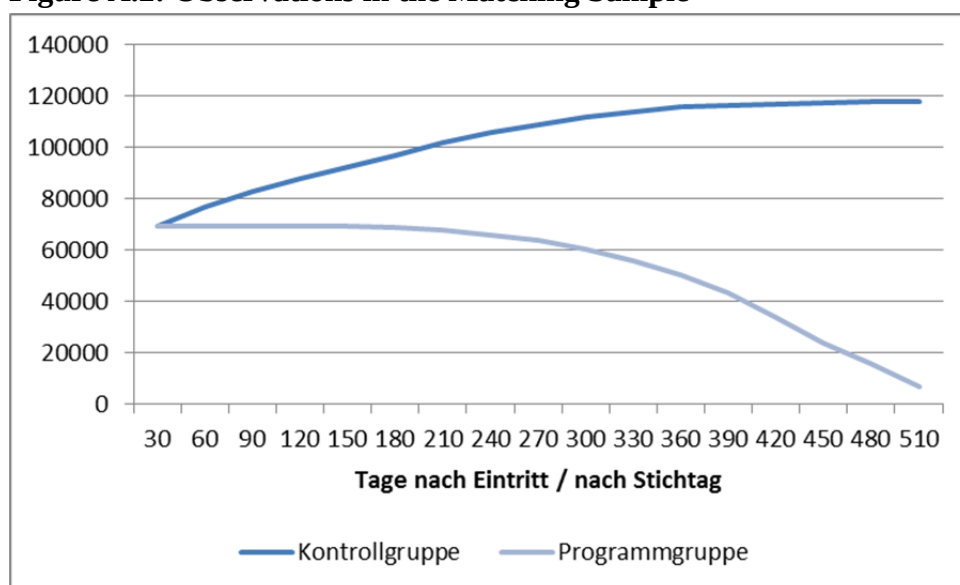
8. Appendix

Figure A.1: Outcome Variables in the Matching Sample over Time



Source: Own calculations based on samples of the Integrated Employment Biographies.

Figure A.2: Observations in the Matching Sample



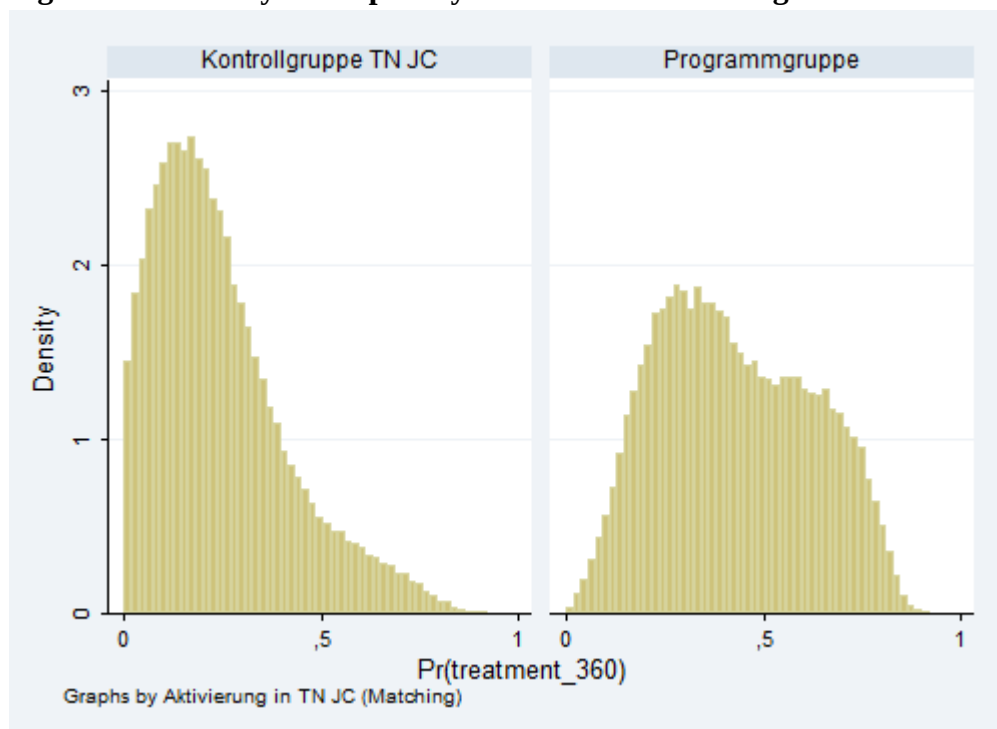
Source: Own calculations based on samples of the Integrated Employment Biographies.

Table A.1: First Stage and Balancing Test Summary Statistics for the Matching Estimation

Probit Estimation	30 Days	180 Days	360 Days	510 Days
Observations	173081	171716	155466	115211
Chi ² Test	29050.74	29002.21	25185.23	6707.55
Pseudo R ²	0.15	0.15	0.16	0.17
Balancing Tests	Pseudo R ²	Likelihood	Chi ² Test p value	
Unmatched	0.16	30025.83	0.000	
Matched	0.02	1955.57	0.000	

Source: Own calculations based on samples of the Integrated Employment Biographies.

Figure A.3: Density of Propensity Scores in the Matching Estimation



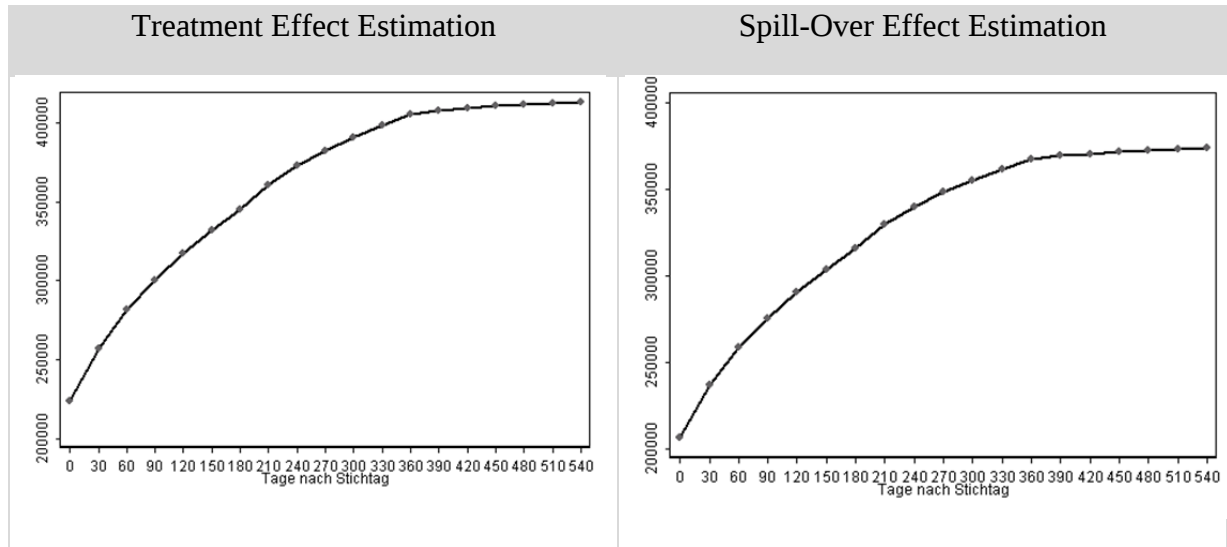
Source: Own calculations based on samples of the Integrated Employment Biographies.

Table A.2: Common Support in the Matching Estimation

Tage nach Eintritt	Treatet		Untreatet
	Off Support	On Support	On Support
30	4.315	39.233	64.067
60	4.394	42.578	70.370
90	4.412	45.566	75.405
120	4.415	48.391	79.673
150	4.440	50.790	83.418
180	4.409	52.871	87.127
210	4.365	54.508	91.245
240	4.240	54.805	94.393
270	4.166	54.194	96.771
300	3.849	51.563	98.798
330	3.617	47.931	100.482
360	3.263	43.310	101.903
390	3.003	37.451	102.630
420	2.566	29.056	103.081
450	2.079	20.356	103.390
480	1.436	13.399	103.637
510	877	5.585	103.823

Source: Own calculations based on samples of the Integrated Employment Biographies.

Figure A.4: Observations in the DiD Estimation



Source: Own calculations based on samples of the Integrated Employment Biographies.